



Research Publication No. 2004-09
November 2004

Why DRM Should Be Cause for Concern: An Economic and Legal Analysis of the Effect of Digital Technology on the Music Industry

By Paul Petrick

This paper is part of the Berkman Center's
Research Publication Series for Berkman Affiliates.
The paper is also available via
the Social Science Research Network Electronic Paper Collection:
http://papers.ssrn.com/abstract_id=XXXXXX

WHY DRM SHOULD BE CAUSE FOR CONCERN:
AN ECONOMIC AND LEGAL ANALYSIS
OF THE EFFECT OF DIGITAL TECHNOLOGY ON THE MUSIC INDUSTRY

By Paul Petrick*

ABSTRACT

In response to piracy and online file trading, the music industry has begun to adopt technological measures, often referred to as digital rights management (DRM), to control the sale and distribution of music over the Internet. Previous economic analysis on the impact of DRM implementation has been overly simplistic. A careful analysis of copyright law and the microeconomic principles governing the music industry demonstrates that commentators have failed to account for factors relevant to the measure of social welfare within the music industry. This paper develops a more refined economic model that is better suited to accurately assessing how legal or technological changes like DRM will affect the music industry.

Utilizing a refined economic model, the analysis suggests that the economic effects of implementing DRM technology are generally negative, albeit uncertain. While DRM implementation may inhibit piracy, facilitate price discrimination, and lower transactional costs, it will likely decrease social welfare by raising barriers to entry and exacerbating a number of existing market failures. Specifically, DRM implementation may facilitate the extension of monopoly pricing, decrease the amount of information available to potential music consumers, diminish the number of positive externalities, and raise artistic and informational barriers to entry into certain genres of music.

Keywords: DRM, Music, Technology, Copyright

* Founder and co-owner, Headtunes Recordings. J.D., 2004, Harvard Law School. This Article was inspired by the Author's experience in the music industry, including his work as a composer, producer, DJ, and club promoter. It is released as a working paper under the Berkman Center's Publication Series for Berkman Affiliates.

WHY DRM SHOULD BE CAUSE FOR CONCERN:
AN ECONOMIC AND LEGAL ANALYSIS
OF THE EFFECT OF DIGITAL TECHNOLOGY ON THE MUSIC
INDUSTRY

By Paul Petrick

Table of Contents

I. Introduction	4
II. The State of Technology	5
A. Digital Technology	5
B. Digital Right Management.....	7
III. The Legal Setting.....	9
A. Exclusive Rights under Copyright.....	9
B. Exceptions to the Exclusive Rights under Copyright.....	10
C. Structural Limitations to the Control over Copyrighted Works.....	13
IV. Refining the Economics of Copyright	14
V. Economic Effects of DRM.....	25
VI. Conclusion	30

Why DRM Should Be Cause for Concern: An Economic and Legal Analysis of the Effect of Digital Technology on the Music Industry

by Paul Petrick

I. Introduction

Digital technology has upset the degree of control copyright holders have historically maintained over the use of and access to music. Computers, and related digital devices and components, as well as peer-to-peer (P2P) networking software, provide the populace with relatively inexpensive and effective means to copy and distribute music in digital formats without legal authorization. While the effects of such technology on the music industry as a whole are still debatable, as digital technologies become more ubiquitous, piracy and online file sharing continue to pose an imminent threat to the ability of artists to derive a sustainable profit from the creation of music.

This purpose of this paper is to determine the effects of utilizing one proposed solution to piracy and online file sharing of music, namely the use of digital rights management (DRM) technology. At its core, DRM technology allows a copyright holder to define a set of rules attached to a work in a digital format that control consumer access, use, and manipulation of that work.¹ Various rules control the ability to copy, archive, and distribute digital works.² In determining the effects of DRM, the analysis will rest firmly on the utilitarian principle of maximizing overall social welfare. The hope is to suggest whether legislative proposals in the realm of copyright and contract law should support or impede music industry investment in the development of DRM technologies.

Though the analysis is not absolutely determinate in its conclusions as a theoretical matter, it does suggest that a number of factors contributing to social welfare have been overlooked or undervalued in the analysis until now. Most notably, previous analysis of the DRM solution has failed to accurately account for losses in utility attributable to increased creation costs caused by DRM implementation.

A couple points should be noted at the outset. First, this paper will focus its analysis solely on music, excluding all other types of copyrighted works capable of being produced and distributed in digital formats. Many of the commentators on DRM have based their analysis on improper or overly simplistic assumptions about how art generally, and music specifically, is created. Focusing on music exclusively will provide concrete examples that refute those assumptions. Moreover, the music industry is on the front line of the battle against piracy and online file sharing, the very behavior DRM is intended to stop.

Second, this paper will be rather speculative in its description of the music industry's implementation of DRM technology. Though various methods of DRM implementation currently exist, DRM technology has not yet provided all of the technological protections many, including myself in this paper, claim are possible. However, DRM will remain a sustainable solution only to the degree it delivers on those technological promises. Thus, in looking toward

¹ GartnerG2 and The Berkman Center for Internet & Society at Harvard Law School, Copyright and Digital Media in a Post-Napster World, August 2003, at <http://cyber.law.harvard.edu/home/uploads/254/2003-05.pdf> (last visited May 12, 2004).

² Id.

the DRM solution's future effects, one necessarily must speculate as to the nature of its implementation.

Part II of this paper will explore the various technologies utilized in the new digital era. Section A will explore the rise of digital technologies, especially P2P networks and sampling, and how they have revolutionized the way music is created, published, and distributed. Analysis will focus on both the advantages and disadvantages digital technology offers to consumers and artists. Section B will explain the mechanics of DRM technology and how that technology proposes to cure the ill effects of digital technologies and P2P networks.

Part III of this paper will examine the current legal structure surrounding artistic creations generally and music specifically. Section A will describe the legal entitlements created by copyright law. In keeping with the utilitarian focus of the paper, it will outline the common economic justifications proffered for the exclusive rights provided by the copyright statute. Section B will describe a few of the exceptions to the exclusive rights described in Section A, again providing the common economic justifications proffered for those exceptions. Section C will examine the structural limitations of control over copyrighted works inherent in the copyright system prior to the development of digital and DRM technology.

Part IV will take a closer look at the economics of music creation. In doing so, it will suggest ways in which commentators have failed to accurately account for factors relevant to the measure of social welfare within the music industry. Applying microeconomic principles, it will suggest a slightly more refined economic model that is better suited to accurately assessing how legal or technological changes like DRM will affect the music industry.

Part V will utilize the economic principles developed in Part IV to determine the effects of DRM on the music industry. Each of the principles will be applied in turn to determine the effects on social welfare. The assessment of DRM will center on whether it incorporates the advantages of digital technology in general while minimizing the costs of its disadvantages.

Part VI will rely on the analysis of Part V to suggest that total social welfare from the creation and consumption of music is likely to decrease under a system that utilizes DRM technology. Specifically, this part will argue that DRM technology would likely decrease the overall number and variety of musical works with the effect of decreasing competition.

II. The State of Technology

A. Digital Technology

Digital technology has revolutionized the reproduction of music. Digital technology provides consumers the ability to generate identical copies of digital music files. Compression formats, most notably MP3, allow consumers to compress digitized music into smaller files. "Ripping" software allows consumers copy music from CDs, store the files on their hard drives, and convert the files into the compressed formats. Computer storage space and processor speeds continue to increase at a staggering pace. Digital file reproduction devices, such as CD writers, are becoming increasingly accessible to consumers as they continue to decline in price. These factors combined make copying music nearly costless and instantaneous.

Digital technology has also revolutionized the dissemination of music. Most prominent among the technological developments has been the creation P2P networks. The software known as Napster created both a technological and legal uproar when it introduced the prospect of sharing music files across the Internet. At its core, P2P networks offer users the ability to access the hard drives of other users worldwide merely by installing a piece of software. Once the software is installed, users can search for, copy, and transfer music files, typically in a compressed format such as MP3, stored in particular locations in the hard drives of other users, and vice versa. Napster, in its original form, compiled a list of the files made available by all its users. The compiled list was held by the company itself in a centralized database. Later versions of P2P networks, most notably KaZaA and Grokstar (collectively, the “FastTrack” providers), neither compile nor control a list in a centralized database. These networks are essentially self-organizing, meaning individual users’ computers compile their own limited index of files available for downloading. Instead of accessing a single, centralized database, users searching for a file access multiple individual computers using those computers’ indexes to determine where to find the files. Thus, FastTrack companies take no prominent part in the actual trading of files, as they control merely the distribution of the software, not its use.

Additionally, digital technology has revolutionized the creation of music. At the heart of that revolution is the digital audio workstation (DAW) and a method for reusing recorded audio, called sampling, which is increasingly used in the creation of music. Along with the ability to make perfect digital copies of recorded music has come the development of hardware to convert music from an analog to a digital format. Once the music is in a digital format, a DAW allows one to import, cut, copy, layer and manipulate (through effects processors) digital copies of the music, analogous to how word processors allow one to cut, copy, and manipulate digital representations of words or passages of words. Sampling is merely the use and manipulation of digital copies of a recorded musical work in the creation of a new musical work. Samples can take the form of passages of an entire band’s recording of a song, passages from just a single instrument, or a single note (“one-shot”) of a single instrument. An artist may use samples of instruments she has recorded herself or samples from other artists’ recordings.

The effect of DAWs on the creation of music cannot be understated. Sampling has become a ubiquitous practice in both the professional and hobbyist music studios. Hip-hop and dance music have come to rely heavily on sampling to define their respective genres, but even pop, rock, country, and jazz songs often use samples of one sort or another. Most studios today contain entire libraries of sampled instruments in the form of recorded passages or one-shots. As the price of DAWs, like other digital technologies, continues to decline, so does the price of creating musical works, especially music that utilizes or is derived from sampled segments of preexisting musical works. The investment in equipment necessary to create professional grade music has decreased from tens of thousands of dollars to hundreds of dollars in the past ten years (assuming one already owns a personal computer).

The advantages to the music consumer from these digital technologies are clear. Putting aside for the moment whether it is legal, digital technology provides cheap and near instant copying and distribution of music. As a result, the average music consumer (at least those with computers and access to the Internet) can access exponentially more music than ever before. Of particular importance, consumers can create multiple copies for personal use in different contexts. For example, I can create copies of my favorite CD, one for the car, one for my home stereo, and one downloaded onto an MP3 player. Furthermore, P2P networks provide a community where consumers can network with fellow users with tastes for particular types of music similar to their

own, effectively minimizing the search costs for music within a particular genre they will enjoy. Once connected, consumers can sample music far more conveniently than in the past. For example, consumers have typically had to rely on friends, the radio, or brick and mortar stores to provide samples of music. If a consumer enjoyed a particular style of music that was not popular amongst his friends, not played on the radio or was not available in their local store, the consumer effectively had no means by which to find music to their taste. Now a consumer merely needs to connect to a network, find another user with similar tastes, then download and listen to music from that user's hard drive. Music that the consumer finds disagreeable can be instantly discarded. Finally, and most profoundly, digital technology spurs consumer participation in the creation of music. Utilizing DAW software, otherwise non-artistic consumers can layer samples of downloaded or copied music to create new works.³

The advantages to the music industry and artists are equally impressive. First, cheap copying and distribution, at least on a theoretical level, reduce the cost of providing music to the public, increasing the number of potential consumers. Copying and distributing music files on the Internet is dramatically cheaper than producing a CD and shipping the contents to brick and mortar stores. In addition, P2P networks potentially provide an enormous marketing tool. Some technologies companies have begun develop software that can track downloading habits of users on P2P networks and aggregate that data by region or city. Such aggregated data can provide the music industry information on the particular artists or styles of music that are popular in a given region. That information can in turn be used to develop targeted marketing plans. Furthermore, P2P networks might be used to market test new artists. Additionally, P2P networks provide one of the components necessary to enable the music industry to better price discriminate, increasing the number of potential consumers and increasing overall economic efficiency (see discussion in Part IV for details). Instead of being forced to offer complete albums, online distribution offers the potential for sales of individual songs, or even individual listens of songs or albums. Finally, as previously discussed, DAWs and sampling are providing increasingly inexpensive means for the creation of music.

While there don't appear to be any disadvantages to the consumer from digital technology (assuming a lack of legal repercussions for the use of the technology), the disadvantages to the music industry are potentially catastrophic. Just as inexpensive copying and distribution of music lowers the costs of providing music to consumers, consumers' ability to cheaply copy and distribute music threatens one of the music industry's primary revenue streams. Once music is sold to any individual consumer, that consumer can make the music available to possibly millions of other potential consumers on P2P networks roughly free of charge. This threatens the music industry's incentive to invest in the creation of works in the first instance. Without the ability to recoup the investment spent on finding artists, recording and then marketing that music, the entire music industry has the potential to implode.

B. Digital Right Management

Partially to offset the disadvantages described above, the copyright industries, including the music industry, have recruited technology companies to provide a means to safeguard against consumer copying and distribution of music. The result is what has been termed digital rights management. DRM technology can be defined generally as "secure packaging and delivery software designed to prevent purchasers and third parties from making unauthorized uses of

³ See, e.g., <http://www.ritmic.com/hitmania/especiales/bootleg/losmejores.html> (last visited May 12, 2004)

digital works.”⁴ The technology, still in development, will ultimately attempt to control, monitor, and meter almost any conceivable use of digitized music.⁵ Most importantly for consumers, DRM would provide the music industry with the means to charge consumers for each use of a music file and prohibit a variety of uses, including piracy and free file sharing.

Commentators have proposed three possible methods by which to implement DRM technology. The first and most direct method would be through the use of computer code. Under this method (hereinafter referred to as the “code-only” method), certain allowable uses would be programmed directly onto the rule set that controls access to a digital file.⁶ Whether the user is charged for a particular use is dependent simply on whether the underlying code requires the user to pay. An example, though certainly not the only possible example, of this type of implementation is the iTunes software. After charging to download a song to a user’s computer, the iTunes software allows the user to make a certain number of copies of playlists and move the song file, with certain limitations, to other specified devices, like MP3 players.⁷ However, this degree of control hardly represents the entire landscape of computer code facilitated control. The paper will assume DRM will be capable of automatically charging for every listen, copy, distribution, manipulation (such as sampling), or other possible use of every song in digital format, regardless of whether it is bought or retrieved online or through a CD or other storage medium. These charges, often referred to as micropayments, would occur automatically, without the user having to take any action other than engaging in the desired use.

The second method for controlling consumer use of digitized music through DRM technology would be through the use of key access. Under this scheme, DRM technology would utilize an external, human decision-maker.⁸ Users would apply for digital keys to access files.⁹ The human decision-maker could then judge on a case-by-case basis whether to allow and whether to charge for the particular use requested by the user. As two commentators have noted, this method “build[s] in judgment capabilities that cannot practically be emulated by technical defaults.”¹⁰ The human decision-maker might be the copyright holder or a neutral third party, for example, a government body.¹¹

The third, and most likely, method would combine the first two methods. Under this scheme (hereinafter referred to as the “code-plus” method), certain allowable uses, and forced charging of the user for, or blocking of those uses not categorized as allowable, would be coded directly onto the file.¹² If the user wished to use a file in a manner not allowed by the code, the user could then apply for an access key.¹³ The obvious advantage to such a system over the other methods is the quick automation of common uses with the added ability to inject a human intelligence into the decision-making process.

⁴ Dan L. Burk & Julie E. Cohen, Fair Use Infrastructure For Rights Management Systems, 15 Harv. J.L. & Tech. 41, 48 (Fall, 2001).

⁵ Id.

⁶ Id. at 55.

⁷ <http://www.apple.com/itunes/> (last visited May 12, 2004).

⁸ Burk & Cohen, supra note 4 at 58.

⁹ Id.

¹⁰ Id.

¹¹ Id.

¹² See id. at 65.

¹³ See id. at 65.

As an initial practical matter, only the code-only and code-plus options seem remotely viable solutions for the music industry. One can easily imagine the tremendous manpower it would take for decision-makers to determine whether every use of every music file was deserving of an access key. The labor force necessary to employ for such an undertaking would far outweigh any potential benefits in combating piracy and music file sharing. Thus, the remainder of this paper will concern itself strictly with an analysis of the code-only and code-plus methods.

III. The Legal Setting

A. Exclusive Rights under Copyright

Under the United States Constitution, Congress is empowered “[t]o promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries.”¹⁴ Copyright law provides the legal mechanism for the promotion of music.¹⁵ Specifically, copyright law grants the authors and their assigns certain exclusive rights, including the rights of reproduction, production of derivative works, distribution, and public performance an authored work.¹⁶ The exclusive rights are granted for a term of 70 years after the death of the author.¹⁷

Before continuing, some definitions are probably in order. In the context of copyright law, the definition of reproduction largely mirrors common conceptions of a copy. Specifically, a reproduction of a sound recording is a duplication of the recording through direct or indirect recapture of the sounds fixed in that recording.¹⁸ However, as we will see in moment, the doctrine covering infringement of the exclusive right to copy subscribes to a broader definition in practice. The definition of a derivative work generally is a work “based upon one or more preexisting works.”¹⁹ With respect to sound recordings, derivative works are works “in which the actual sounds fixed in the sound recording are rearranged, remixed, or otherwise altered in sequence or quality.”²⁰ Distribution is defined as the “sale or other transfer of ownership, or . . . rental, lease, or lending” of work.²¹ Finally, public performance of a recording is a digital

¹⁴ U.S. Const. art. I, § 8, cl. 8.

¹⁵ See 17 U.S.C. § 102 (1996) (listing the types of works protected by copyright)

¹⁶ Section 106 of the Copyright Act provides:

[T]he owner of copyright under this title has the exclusive rights to do and to authorize any of the following:

to reproduce the copyrighted work in copies of phonorecords;

to prepare derivative works based upon the copyrighted works;

to distribute copies or phonorecords of the copyrighted work to the public by sale or other transfer of ownership, or by rental, lease, or lending;

in the case of literary, musical, dramatic, and choreographic works, pantomimes, and motion pictures and other audiovisual works, to perform the copyrighted work publicly;

in the case of literary, musical, dramatic, and choreographic works, pantomimes, and pictorial, graphic, or sculptural works, including the individual images of the motion picture or other audiovisual work, to display the copyrighted work publicly; and

in the case of sound recordings, to perform the copyrighted work publicly by means of a digital audio transmission.

17 U.S.C. § 106.

¹⁷ 17 U.S.C. § 302.

¹⁸ 17 U.S.C. § 114.

¹⁹ 17 U.S.C. § 101.

²⁰ 17 U.S.C. § 114.

²¹ 17 U.S.C. § 106.

audio transmission, commonly referred to as webcast.²² Notably, playing a recording over the airwaves does not qualify as a public performance of the recording.

Copyright infringement may occur under a number of different contexts. Generally, infringement of the right to copy has occurred when there was actual unauthorized copying of part or all of a protected work and the copying “went so far as to constitute improper appropriation” of, or the new work was substantially similar to, the protected work.²³ For example, ripping and burning a copy of a CD for commercial use would qualify as infringement. Infringement of the right to create derivative works essentially uses the same test. Thus, an infringement of the right to create derivative works is often equated with infringement of the right to copy, especially as the law relates to sound recordings. One example of infringement of the right to create derivative works is the unauthorized use of samples, although the length of the sample necessary to qualify as infringement remains unclear.²⁴ Infringement of the right to distribute occurs when a work has been distributed to others without authorization. Digital distribution through P2P networks is one example of this form of infringement (as well as infringement of the right to copy).²⁵ Finally, infringement of the right to publicly perform a recording generally occurs when webcasters stream music without paying statutory licensing fees. Importantly, each of the described instances of infringement is subject to one or more of the exceptions outlined in the next section.

The economic justification typically supplied for the exclusive rights granted to authors necessitates a discussion of public goods. Copyrighted works, music included, are often characterized as public goods. Public goods are defined by two characteristics. First, public goods are nonrivalrous, meaning they can be used and enjoyed by an infinite number of persons without ever being used up.²⁶ Second, public goods are nonexclusive, meaning that it is difficult to prevent people from accessing, and thus enjoying, the goods.²⁷ Together, these characteristics create a disincentive to create works by enabling the public to access and enjoy a work without necessarily compensating the author.²⁸ The exclusive rights granted by copyright attempt to counteract the negative impact of the public goods phenomena. As Wendy Gordon explains, “if the creators of intellectual productions were given no rights to control the use made of their works, they might receive few revenues and thus would lack an appropriate level of incentive to create.”²⁹ As a result, “[f]ewer resources would be devoted to intellectual productions than their social merit would warrant.”³⁰

B. Exceptions to the Exclusive Rights under Copyright

Though there are a number of statutory exceptions to the exclusive rights granted by copyright law, this paper will primarily discuss the two primary exceptions that affect musical recordings – fair use and first sale – and then quickly describe a few other exceptions of more minor importance. The first exception to the exclusive rights provided to authors of copyrighted works

²² 17 U.S.C. § 114.

²³ See *Arnstein v. Porter*, 154 F.2d 464 (1946).

²⁴ See *Newton v. Diamond*, 349 F.3d 591 (9th Cir. 2003) (providing exception for de minimis use).

²⁵ See *A&M Records v. Napster*, 239 F.3d 1004 (9th Cir. 2001).

²⁶ William V. Fisher, III, *Reconstructing the Fair Use Doctrine*, 101 Harv. L. Rev. 1661 (1988).

²⁷ *Id.*

²⁸ *Id.*

²⁹ Wendy J. Gordon, *Fair Use as Market Failure: A Structural and Economic Analysis of the Betamax Case and Its Predecessors*, 82 Colum. L. Rev. 1600, 1611 (1982).

³⁰ *Id.* at 1610.

is fair use. Under the copyright statute, unauthorized use of a copyrighted work for “criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research” may provide a defense against a claim of copyright infringement.³¹ However, the actions enumerated above do not describe the range of possible uses that fall under the fair use defense. Fair use is decided on a case-by-case basis. To that end, the copyright statute provides a four factor test to determine whether a particular use is considered fair. The factors are: 1) the purpose and character of the use, including whether the use is commercial or for non-profit educational purposes; 2) the nature of the copyrighted work itself, in which the court typically evaluates whether the work is factual, scientific, or artistic in nature; 3) the quantitative amount and substantiality, interpreted as the qualitative amount, of the work that is copied; 4) the effect of the use upon the potential market for or value of the work.³² In amending the copyright statute, Congress stated that the four factor test was “intended to restate the present judicial doctrine of fair use, not to change, narrow, or enlarge it in any way.”³³

Commentators have proposed two separate economic justifications for the creation of the fair use doctrine. Both are based on market failures, that is, the inability of the market to efficiently provide copyrighted goods.

First, one commentator has suggested that fair use may serve as a device for correcting the market failure that occurs when the transaction costs for negotiating a license for a copyrighted work are so high as to prohibit or deter a valuable use of that work.³⁴ For example, a music teacher at a local community center wishing to photocopy sheet music for use in a classroom discussion would undoubtedly find it too troublesome to contact the publishers and negotiate licenses for each photocopy. Even if the teacher is willing to pay for the licenses and can pass the costs onto his paying students, without fair use, the added cost in contacting and negotiating with a number of publishers will deter the teacher from utilizing the books for an otherwise valuable purpose.

Second, a number of other commentators have suggested that fair use may serve as a device for correcting another type of market failure that occurs when the value of a particular use of a copyrighted work cannot be fully internalized.³⁵ This justification centers on two related but separate concepts. In one sense, commentary, criticism, parody, and other unauthorized uses serve a socially valuable purpose by fostering public debate and contributing to an informed populace.³⁶ However, the diffuse value it provides to the public at large cannot be incorporated into the price of the work because the potential consumer will be unwilling to pay for the value that accrues to the larger public.³⁷ In the realm of music, good examples of this concept are hard to come by. However, songs which incorporate samples of other popular songs, technically

³¹ 17 U.S.C. § 107.

³² *Id.*

³³ H.R. Rep. No. 94-1476, at 66 (1976), reprinted at 1976 U.S.C.C.A.N. 5659, 5680; S.Rep. No. 94-473, 62 (1975).

³⁴ Gordon, *supra* note 29, at 1600.

³⁵ See Robert P. Merges, Are You Making Fun of Me?: Notes on Market Failure and the Parody Defense in Copyright, 21 AIPLA Q.J. 305 (1993).

³⁶ See Julie E. Cohen, Lochner in Cyberspace: The New Economic Orthodoxy of “Rights Management”, 97 Mich. L. Rev. 462 (1998); Lydia Pallas Loren, Redefining the Market Failure Approach to Fair Use in an Era of Copyright Permission Systems, 5 J. Intell. Prop. L. 1 (1997).

³⁷ *Id.*

derivative works of the songs sampled, are often cited as an example of fair use.³⁸ Courts have sometimes called such works parody, implicitly or explicitly arguing the songs serve as a means to critique or comment on the work sampled. In another sense, some private, noncommercial copying and distributing of copyrighted works may generate value to the public by inspiring the creation of new works and contributing to the free flow of ideas.³⁹ Again, examples in the music context are difficult to pin down. For now, suffice it to say that music has potential positive externalities that a copyright holder cannot assess nor recoup when selling it; fair use provides a means to subsidize uses that create sizeable value.

The second relevant exception to the exclusive rights granted by copyright is the first sale doctrine. Under the copyright statute, copyright holders are entitled to the right of control only over the first public distribution of a particular copy of the work.⁴⁰ For example, once a record company sells a CD to a retailer and a consumer buys it, the company cannot control the ability of the consumer to resell the CD or demand that the consumer charge a specific price for the used CD. However, there is an exception to this exception in the context of sound recordings. Sound recordings may not be rented, leased, or lent for commercial purposes without authorization.⁴¹

Despite the foregoing, the first sale doctrine as it applies to digital works is still somewhat unclear. As a preliminary matter, the first sale doctrine does not explicitly distinguish between analog (e.g. tape) and digital recordings. The reason for the lack of clarity stems in part from one court's decision to classify copies of digital works in RAM as reproductions.⁴² For example, suppose I have legally paid for and downloaded a song in digital format. Now suppose I dislike or have grown bored of the song and wish to resell it. The problem with reselling it through electronic transmission, i.e. by way of the Internet, is that the actual file itself is not transmitted. Rather, a copy is produced on the buyer's computer and the original file is left intact. Moreover, even if the buyer deletes their own copy, multiple copies are made in RAM – a temporary storage space on computers – on both the buyer's and the seller's computers during the process of transmission. Thus, resale of digital works through networks implicates the copyright holder's exclusive right to reproduction, to which the first sale doctrine does not apply.

Historically, the justification for the first sale doctrine has been to prohibit restrictions on the alienation of tangible property, just as in the traditional property context. However, recent utilitarian justifications have centered on the effects it has on prices. The ability of retailers and consumers to determine the price they sell works promotes competition in the retail (primary) market and promotes the development of secondary markets.⁴³ As a result of competition in these markets, prices are pushed downward. Decreased prices increase access to works.

³⁸ See, e.g., Burk & Cohen, *supra* note 4 at 44 (proposing that the Supreme Court's decision, in *Cambell v. Acuff-Rose Music, Inc.*, that a rap group's use of a rock sample, despite having been denied a license for its use, qualified as fair use was an example of the court recognizing the critical value in the work).

³⁹ Yochai Benkler, *Free as the Air to Common Use: First Amendment Constraints on the Enclosure of the Public Domain*, 74 N.Y.U. L. Rev. 354 (1999); Julie E. Cohen, *Copyright and the Perfect Curve*, 53 Vand. L. Rev. 1799 (2000); Neil Weinstock Netanel, *Copyright and a Democratic Civil Society*, 106 Yale L.J. 283 (1996).

⁴⁰ 17 U.S.C. § 109.

⁴¹ *Id.*

⁴² *MAI Systems Corp. v. Peak Computer, Inc.*, 991 F.2d 511 (9th Cir. 1993), cert. dismissed, 114 S. Ct. 671 (1994).

⁴³ See, e.g., R. Anthony Reese, *The First Sale Doctrine in the Era of Digital Networks*, 44 B.C. L. Rev. 577.

A few other exceptions are noteworthy. The Audio Home Recording Act, which amended the copyright statute, provides an exception for the noncommercial reproduction of recordings by consumers for personal use.⁴⁴ Thus, for example, a consumer may copy a CD, use a sample to create a derivative work, and send digital copies from one computer to another, as long as consumer alone uses the resulting work for noncommercial purposes. Additionally, libraries or archives may make three copies of a recording or may create a replacement copy if one of the copies is damaged, lost, or stolen.⁴⁵ Finally, webcasters generally need not pay statutory licensing fees if the webcasters broadcast under a non-subscription service licensed by the FCC.

Economic justifications for these exceptions are unclear. In some sense, these may be seen as limiting the scope of exclusive rights merely to maintain some general level of access. However, this is not technically a strict utilitarian argument, but rather a distributive justification. With that said, allowing consumers to make multiple copies for personal use or use by libraries and non-subscription webcasters may create positive externalities large enough to merit subsidization.

C. Structural Limitations to the Control over Copyrighted Works

Contrary to claims made by the music industry, piracy and the sharing of music has long been the norm amongst many music consumers and artists, though perhaps on a far smaller scale than what occurs today. Before the introduction of the CD, consumers often made unauthorized tape copies of complete albums for the purpose of distributing them to friends. Similarly, the mix tape, a compilation of copied songs from various albums, was ubiquitous as well. Additionally, consumers often recorded music from the radio or at concerts onto tapes without authorization. At an artistic level, copying has an even longer tradition. Musicians have long copied passages, not of the sound recording, but of the composition of other artists. Jazz musicians are notorious for musically “quoting” melody lines of songs or other musicians’ solos – arguably, in copyright jargon, reproducing a work or producing a derivative work – when composing or soloing themselves.

Though each of these practices was, and still are, arguably illegal under the law, they have long been tolerated. There are a few plausible reasons. First, before the introduction and mass acceptance of consumer CD writing devices, creating perfect copies of songs was generally cost-prohibitive. Moreover, degradation of sound quality resulted from each subsequent copy of a work on many consumer-level formats. Thus, there was, in a sense, a natural barrier to continued copying. This natural barrier ensured that tape to tape, radio to tape, and concert to tape recordings did not infringe dramatically on the primary market for sound recordings. Second, and still applicable today, consumers find additional value in owning the packaging of a sound recording. Many consumers reap pleasure from owning the track listing, graphic work, lyrics, and photos that accompany a retail version of a sound recording. Those who would state otherwise, sometimes even the music industry itself, overlook the obvious evidence illuminated by the industry practice of providing such additions to the sound recording. Third, prosecuting such behavior was largely impractical. Even if artists and record companies were willing to bear the cost of litigating minor infractions for the sake of deterrence, as recent lawsuits against file sharers suggest they might have been, such deterrence was probably outweighed by the costs associated with identifying those engaged in the illegal practices.

⁴⁴ 17 U.S.C. § 1008.

⁴⁵ 17 U.S.C. § 108.

Given the tradition of pirating and sharing music, it should not be surprising that as music took form in digital formats, consumers continued to behave in a similar manner. As previously explained, digital technologies merely decreased the cost of copying and sharing music. Thus, the common complaint that an entire generation of children and young adults today is devoid of the moral integrity or understanding of the law ignores the manner in which tradition has shaped behavior. As the next section will suggest, such complaints also serve to obfuscate the social welfare benefits inhering in such behavior.

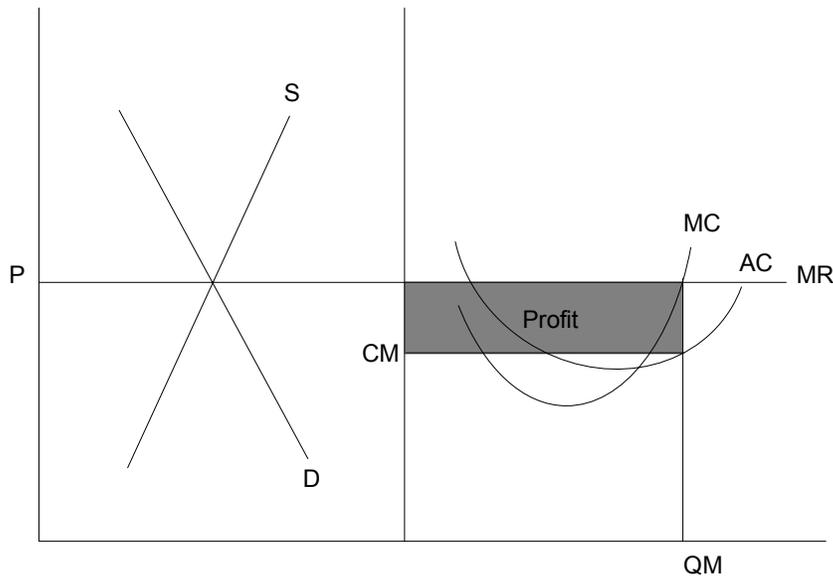
IV. Refining the Economics of Copyright

As is often explained, the exclusive rights granted by copyright law engender a compromise between the interests of the public and the interests of individual authors. As previously outlined, the exclusive rights promote the progress of the arts by creating incentives for artists to produce works. Conversely, the limited scope of the exclusive rights (rights of reproduction, derivative works, distribution, and public performance for a limited duration) ensure the public continues to gain access to those works. Additionally, commentators justify exceptions as curing market failures (fair use) or increasing price competition (first sale), thus increasing overall access and social welfare. These characteristics have given rise to the notion of the access/incentive tradeoff in copyright law.

In order to better qualify the assertions above, let us venture into the world of microeconomics. The benchmark case for examining markets is the generalized case of a market with perfect competition. Perfect competition makes a number of assumptions. First, all firms in the market sell an identical product. Second, all firms are price takers, that is, must charge the price the market as a whole produces. Third, all firms have relatively small market share. Fourth, buyers are well informed about the nature of product offered and the prices at which each firm offers the product. Fifth, the industry is characterized by freedom of entry to and exit from the market.

Figure 1 provides a graphical representation of a perfectly competitive market for a rivalrous and exclusive good from the point of view of the industry and a (profitable) producer within that industry. The x axis and y axes represent quantity and units of value (dollar or the like), respectively. In the left box are the supply and demand curves for the industry in general. S, the industry-wide supply, represents the quantity of a good that the aggregation of all producers in an industry for that good would be willing to produce at each price. D, the industry-wide demand, represents the total quantity of a good that the aggregation of all potential consumers of that good would be willing to buy at each price. P represents the equilibrium price at which consumers are willing to buy the product and producers must sell. The shaded area represents the surplus created by the industry. Surplus is the total value that consumers obtain from consuming the good minus the total costs of producing the good. In the right box, MC represents the marginal cost to the supplier of producing the good, that is, the additional cost of producing each additional unit of the good, for each discrete quantity. AC represents the average cost to the supplier of producing the good for each discrete quantity. MR represents the marginal revenue, that is, the additional value the supplier would gain from selling each additional unit of the good, for each discrete quantity. QM represents the quantity of the good the producer should produce in order to maximize profits and CM represents the average cost of producing each good at that quantity. The shaded area represents the profit the producer reaps from selling QM units of a good at price P at an average cost of CM.

Figure 1



A number of characteristics are worth noting. First, that the supply curve (S) slopes upward and the demand curve (D) slopes downward should come as little surprise. As the price of a good increases, additional producers will enter the market and existing producers are likely to produce more of it if they can take advantage of economies of scale. Conversely, as the price of a good increases, consumers are likely to buy less of it because they have better things to spend their money on or they may find cheaper, substitute products.

Second, surplus is maximized at the equilibrium price (P). This is fairly obvious from a geometric standpoint. If the price is increased or decreased, either the total quantity of the good consumed or the total quantity of the good produced will decrease, respectively. The result will be that the total value consumers obtain from consuming the good minus the total costs of producing the good will decrease as well.

Third, in a perfectly competitive market, marginal revenue is represented by a horizontal line. That value (P) is determined by the intersection of the industry-wide supply and demand curves. Specifically, the intersection represents the quantity and price at which producers are willing to sell the good equals the quantity and price at which consumers are willing to buy the good.

Fourth, average cost and marginal cost are generally both “U” shaped and typically intersect. The reason is that a producer incurs both fixed costs, costs that remain the same regardless of the level of production, and variable costs, costs that increase in direct proportion to the level of sales, when producing a good. Variable cost drives the marginal cost curve. The marginal cost initially declines as increases in volume allow the producer to take advantage of economies of scale. Eventually, however, the increase in volume exhausts all economies of scale. Scarce resources and congestion in production facilities then push the variable costs back upward. A combination of variable and fixed costs drives the average cost curve. At relatively low quantities, the initial fixed costs, which are generally quite large, factor significantly into the average cost of each unit. As the quantity increases, the amortization of fixed costs over the

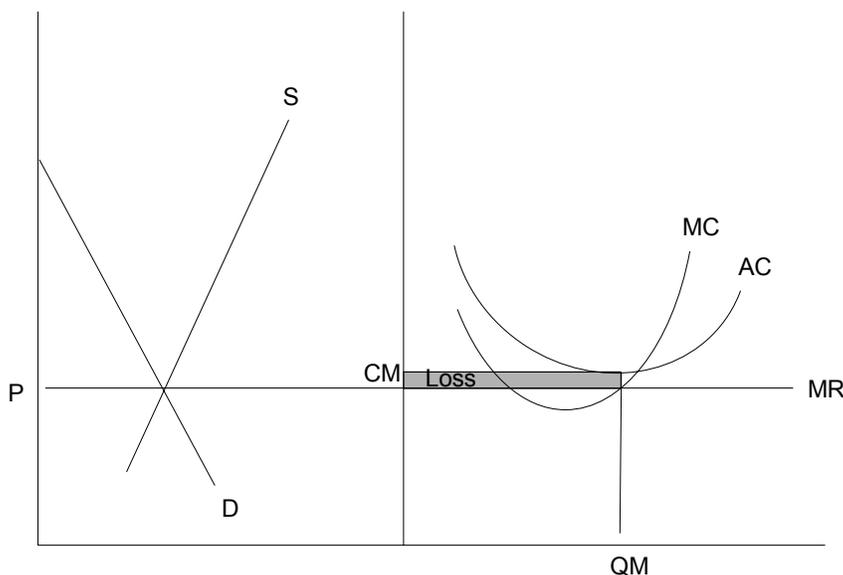
quantity of goods produced combined with economies of scale push the average cost downward. At some volume, variable price becomes the primary factor in the average cost. As the variable prices increase for the reasons previously described, the average cost curve begins to increase as well. At small volumes, the average cost curve remains above the marginal cost curve. However, when goods are rivalrous, at some volume, the average cost curve intersects and drops below the marginal cost curve.

Fifth, profit is maximized at the quantity at which the marginal cost curve and marginal revenue curve intersect. The reason is that at the intersection, the benefit to the producer of producing one more unit of the good no longer exceeds the costs of producing one more unit of the good.

Figure 2 provides a graphical representation of a perfectly competitive market for a rivalrous, but nonexclusive, good from the point of view of the industry and a (formally profitable) producer. Notice that due to the shift in the demand curve (D), the price (P) and the marginal revenue curve (MR) have dropped considerably. Further notice that the shaded area now represents the loss the producer incurs from selling QM units of a good at price P at an average cost of CM.

Why would a formally profitable producer become unprofitable? The reason stems from the fact that the good is nonexcludable. Once the producer sells the good to the public, others are capable of accessing it free of charge. This pushes the price that consumers are willing to pay at each quantity downward, i.e., a number of potential consumers that would forego buying the good, waiting until others bought it and then accessing it for free. The shift in the demand curve would thus lead to a lower equilibrium price (P). The result is that at the quantity where the price consumers are willing to pay for the good intersects the marginal cost of producing that good, the average cost for that good is larger than the price. As the price represents the value consumers place on buying the product, any attempt by the producer to increase the price to make up for the difference will only ensure no one buys the product. Faced with this scenario, a producer will forego producing the good in the first place.

Figure 2



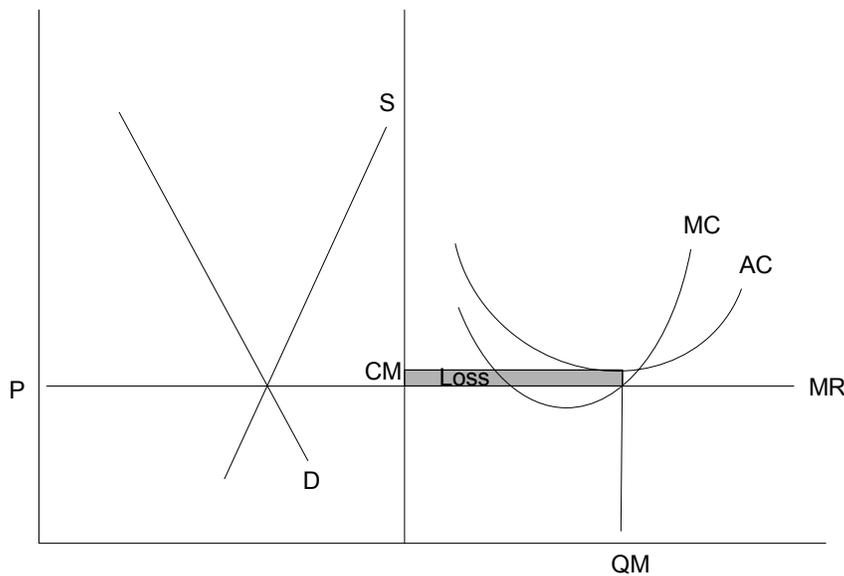
Now let us consider the specific example of sound recordings. Suppose that a record company wishes to create and sell an album.⁴⁶ First, the company must invest in the recording process, including the costs of equipment use and pay for the artists, engineers, and musical producer.⁴⁷ Additionally, it must market the band. It must also pay for photographers, designers, and copyists to create the package layout. Finally the record label must provide some sort of administrative support. Together these represent the fixed costs of creating the album. Now consider the remaining costs. The record company must pay for a tangible medium (tapes, CDs, or hard drive space) on which the product will be distributed, the cost of copying the recording to that tangible medium, the printing costs for the packaging, the distribution costs and other types of overhead. Together these represent the variable costs of creating the album, as the costs are dependent on the number of units produced.

For the moment, consider a world in which copyright and DRM do not exist. Figure 3 is the graphical representation of the competitive market for the album from the point of view of the industry and record company. Though the ultimate outcome in the example will be the same as in figure 2 – the producer, in this case record company, will not be profitable – the reasons are slightly different. One might object that there cannot be a market of multiple producers. Indeed, initially this would be the case. The record company could choose any price and its corresponding quantity along the industry demand curve that would create profit. But soon after the product was made available to the public, it could be easily copied. Opportunistic companies would realize that they need not incur most of the fixed costs associated with bringing the album to market. Specifically, opportunistic companies would not need to invest in the recording process. Additionally, the marketing, designing, and administrative costs would likely be greatly reduced. Provided opportunistic companies had similar variable costs as the record company – in the form of production and distribution capabilities – the average cost curve for the opportunistic companies would closely follow the marginal cost curve of the record company. This would result in an industry-wide supply curve that intersected with the demand curve at a price and quantity that, though profitable for the opportunistic companies, would be unprofitable for the record company. Just as in the case represented by Figure 2, the record company would forego production of the album in the first place.

⁴⁶ Record companies have never invested primarily in producing single song recordings of artists, probably due to the fact that recording a series of songs by any given artist leads to economies of scale. Also, artists themselves rarely bring their own albums to market, also probably due to the fact that a record company provides advantages related to economies of scale.

⁴⁷ This simplification ignores, for the time being, the fact that artists and producers are usually paid royalties which are dependent on the number of units sold.

Figure 3



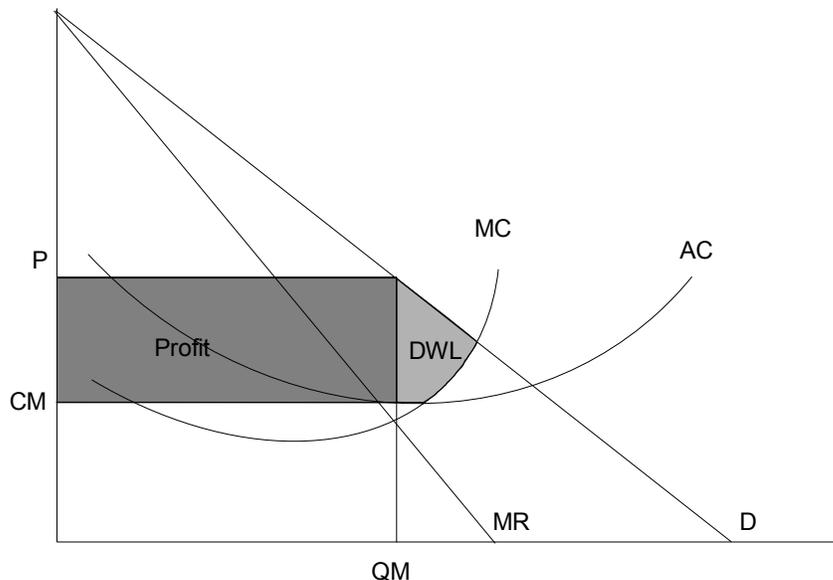
A few points are worth mentioning. Sound recordings, because they are distributed in a tangible medium, are not any more nonexcludable than any other good. A consumer must pay to gain access to them. Second, sound recordings as embodied in a tangible medium are indeed rivalrous. The costs associated with their production is dependent on scarce resources, including recording equipment, the objects on which the sound recording is fixed, labor, and product production, packaging, and distribution facilities. Therefore, the analysis of the record company example suggests that the traditional characterization of copyrighted works, at least sound recordings, as public goods in order to justify copyright law is inaccurate.

Instead, the rationale for copyright's grant of exclusive rights reduces to two related factors. The first is the relative ease with which one can create a copy of the good. The second is the high fixed costs the original producer must incur but that those who create copies do not. Absent copyright law, the producers of sound recordings, like the producers of many other artistic works, are simply cursed with what is termed as the free rider problem. In a world without copyright law, some producers will rely on others to bear a disproportionate cost of bringing the product to market. The result is the underproduction of a socially valuable good.

Now let us consider the album example in the context of copyright law. Figure 4 is the graphical representation of the market for the album from the point of view of the record company. Notice that the graphical representation has changed dramatically. First, there exists a demand curve (D) specific to company which slopes downward. The reason for this is that the exclusive rights granted by copyright ensure that only the original producer of the sound recording, the record company, can sell the album. The demand curve slopes downward for the same reason that the demand curve sloped downward in the previous examples. At lower prices, more consumers would be willing to buy the album, and at higher prices, some consumers would forego buying the album or would find a substitute album at lower price (more on substitutes later). Second, the marginal revenue curve (MR) is no longer flat and now slopes downward beneath the demand curve. The reason is due to the downward sloping nature of the demand curve. In a perfectly

competitive market, individual firms increase their revenue by the same amount, the market price, for each successive unit produced. In the market represented in Figure 4, referred to as a monopoly, the record company must decrease the amount it charges with each successive unit produced. As that decrease in price effects not only the additional unit produced, but the price of all units produced, the marginal revenue decreases as quantity increases. Third, surplus is now represented as it relates to the record company because the record company supply alone represents the total supply in the industry for the album.

Figure 4



The problem with monopoly is that the surplus is not maximized, that is, surplus would increase if the record company produced more units of the album. Recall that surplus is the total value that consumers obtain from consuming the good minus the total costs of producing the good. The quantity and price at which the marginal cost curve and demand curve intersect represents the point at which cost of producing one additional unit of the album will exceed the value consumers obtain from consuming that additional unit of the album. Thus, the price and quantity at that intersection represents the price and quantity that would maximize surplus. However, the record company will neither produce that quantity nor offer the good at that price. The reason is that profit maximization from the company's standpoint occurs at the price and quantity where the marginal revenue curve and the marginal cost curve intersect. The result is that the quantity produced is lower and the price higher than the price and quantity which maximizes surplus. Thus, the lightly shaded triangle represents lost surplus, often referred to as deadweight loss (DWL). The resulting company profits are also much higher than would occur in a competitive market. Those profits are often referred to as being supracompetitive.

Ultimately, useful economic models must represent the entirety of the industry. While Figure 4 describes a market for a single album, it does not accurately represent the market for all albums. Observation of consumer behavior suggests that while there are certainly consumers who make decisions about whether to buy a particular album based purely on the price of that particular album, there are a number of consumers who will also base their decision in part on the price of

other albums. Thus, albums by one artist provide imperfect substitutes for albums by other artists. For example, consider a fan of pop music. Suppose that the fan enjoys Britney Spears, but enjoys Christina Aguilera more. Now suppose that Christina Aguilera's album is priced at \$14 and Britney Spears' album is priced at \$12. Further suppose that at these prices, the fan will buy Christina Aguilera's album in a week's time. Now assume that the price for Christina Aguilera's album increases to \$16 in the course of the week, but that the price of Britney Spears' album remains the same. It is quite possible that the fan will now buy Britney Spears' album. The reason is that the loss in enjoyment the fan will incur from paying a higher price for the Christina Aguilera album is greater than the loss in enjoyment the fan will incur from buying the Britney Spears album instead.

Demand for goods that demonstrate this type of behavior are said to be cross elastic. Specifically cross elasticity of demand for a good is the responsiveness of the demand for the good to a change in the price of another good. The cross elasticity of demand for any particular album will depend on the similarity in style of the available substitutes. Britney Spears and Christina Aguilera are both pop artists. Thus, we would expect the demand for Christina Aguilera albums to respond greatly to price fluctuations in the price of Britney Spears albums. However, we would not expect the demand for Miles Davis albums to respond much, if at all, to price fluctuations in the price of Britney Spears albums.

Now consider a variation of the above example. Consider a different fan that really enjoys country music, but will settle for pop. Now suppose that the fan gets some enjoyment from Britney Spears and enjoys Christina Aguilera only slightly more. Suppose again that Christina Aguilera's album is priced at \$14 and Britney Spears' album is priced at \$12. Further suppose again that at these prices, the fan will buy Christina Aguilera's album in a week's time. Now assume that during the week before the fan plans to buy the Christina Aguilera album, the fan hears songs from a new album by Shania Twain on the radio. After hearing the songs, the fan can't get stop humming them. When the fan finally decides to go buy an album, the fan happily pays \$18 for the Shania Twain album. Why would the fan pay \$18 for the Shania Twain album, but not the significantly cheaper \$14 for the Christina Aguilera album or \$12 for the Britney? The reason is that the total value the fan enjoys from purchasing the Shania Twain album, total enjoyment minus the cost of purchase, is greater than the total value the fan enjoys from purchasing the Christina Aguilera album. As this example demonstrates, the introduction of a work that is diverse in comparison with existing works has the potential to increase the value garnered by an individual consumer beyond the value that would have been garnered by that individual had the work not been introduced into the market.

Now consider a third example. Consider a different fan that enjoys primarily country music. Suppose the fan does not enjoy either Christina Aguilera or Britney Spears enough to buy their albums at the price offered. However, like the previous fan, this fan hears songs from the new Shania Twain album on the radio and loves it. As a result, the fan runs to the store and pays the \$18 for the album. As this example demonstrates, the introduction of a work that is diverse enough in relation to the other works may create new value by creating previously non-existent (or below price level) demand.

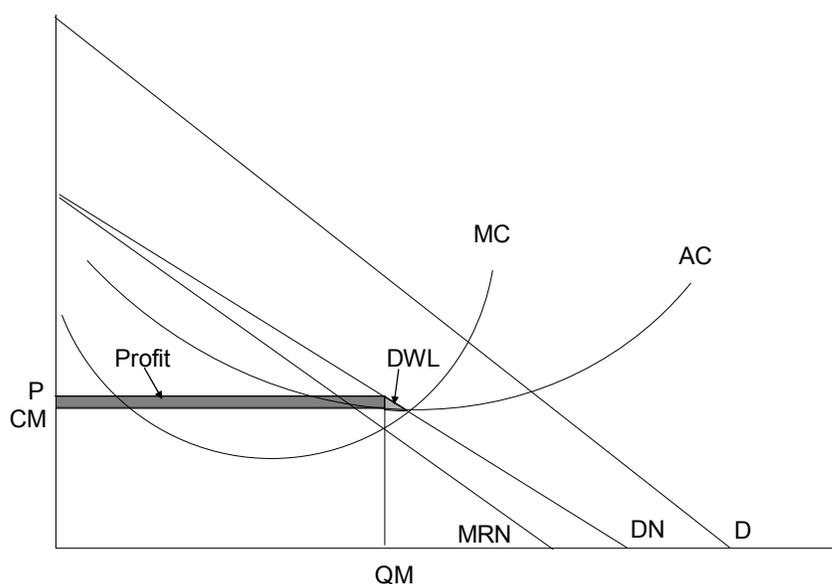
The implications of the examples are profound as they relate to the music industry and producers of sound recordings within the industry. With respect to the industry as a whole, the examples suggest that a shift in the supply curve (for all albums) that increases overall diversity of supply may potentially cause a corresponding shift in the demand curve that increases overall demand.

Thus, total surplus gains from increased supply are greater than the surplus gains resulting from the increased supply and concomitant price decline alone. The total surplus gains also include added surplus from any resultant increased demand. As a general matter, increased musical diversity increases overall social welfare.

The added surplus gains derive from generalizing the results of the second two examples to the industry for albums as a whole. Just as the introduction of the Shania Twain album created or increased value to the fans because it better matched their taste, the introduction of additional albums by different artists into the market either creates or increases the potential value each consumer derives from each purchase because it is more likely consumers will find music best (or at least better) suited to their individual tastes. That the introduction of additional albums into the market necessarily increases diversity (in some mild form, at least) is a rather trivial observation. Because copyright limits reproduction of artistic works, no new album introduced into the market will exactly replicate the existing albums. As consumers respond to the increasing diversity, in the form of musical genres, firms will respond by creating works better targeted to consumers' tastes. As the subsequent albums produced better match existing and potential consumers' tastes, more consumers are likely to purchase albums and the value obtained by each existing consumer from the purchase of an album increases. The result is an overall increase in demand. The increase in demand increases total surplus beyond the surplus gains associated with increased production and the subsequent price decline alone.

There is, however, a limitation to the ability of the market to reap these disproportionate surplus gains. To determine the limiting factor, we must consider the effect of the examples described above at the producer level. Figure 5 is a more refined graphical representation of the market for an album from the point of view of our intrepid record company. Notice that two new curves have been added. The first, DN, represents the new demand curve caused by competition from other albums. The second, MRN, represents the new marginal revenue curve that results from the new demand curve (DN).

Figure 5



A number of characteristics are of importance. First, notice that DN is smaller than D . This is due to the availability of albums that are better suited to the consumers' tastes, just as in the second example. Second, notice that both DN and MRN are flatter than D and MR . This is due to the cross-elasticity of demand for the album with respect to other albums. As the first example demonstrated, demand for the album is dependent in part on the price of other albums. As more albums are introduced into the market, consumers are more likely to find cheaper close substitutes if the price of the album increases. Third, notice that total surplus for the album has decreased, as has deadweight loss.

The last characteristic deserves special attention. Even though the surplus for each album has decreased, remember that total surplus has increased (or at least remains the same). The availability of more albums means that each producer of an album creates and captures a smaller proportion of the total aggregated surplus. Each album divides the total surplus into increasingly smaller fragments. Entry into the market for albums generally and for particular genres specifically will continue as long as new albums can capture part of the total aggregated surplus. Thus, there exists a market limitation on the number and variety of entrants.

Consider further the effects on deadweight loss. Deadweight loss actually decreases in greater proportion than the surplus from the point of view of a producer. This can be seen graphically by the fact that DN and MRN have flattened. As more and more possible substitute albums become available for any particular album, the more responsive the demand for that particular album is to price changes. When DN and MRN are flatter, the difference between the price and quantity at which the record company maximizes profits and the price and quantity at which surplus is maximized becomes smaller in proportion to the surplus resulting from the profit maximizing price. Thus, as the quantity of albums increases, the aggregate deadweight loss decreases and the market begins to more closely resemble a market with perfect competition.

It is worth reiterating the effects of increasing the quantity of albums produced. First, it increases total surplus by increasing the value each consumer realizes from each purchase and increasing overall number of consumers. The increased value stems from the fact that the increased number and concomitant diversity of albums makes it more likely that those existing and potential consumers will find and purchase an album catering to their ideal taste. Second, increasing the amount and diversity of albums lowers total deadweight loss. This stems from the fact that the availability of many possible substitutes to a particular album ensures that the creator of that album engages in less severe monopolistic pricing. These results suggest that policies that stimulate entry into the market, and the concomitant diversity in music, are the best economic prescription for the music industry. Finally, there exists a natural limitation to the surplus gains. Entry into the market for albums generally and the market for albums of a particular genre or style specifically will occur as long as existing albums continue to generate profit. Producers will divide the market into increasingly smaller fragments, pushing prices downward, until there is virtually no profit left to divide amongst new entrants.

The above analysis suggests, then, two general policies for both providing incentives for the creation of and fostering consumer access to music. Those policies are the stimulation of entry into the market and the facilitation of the creation of close substitutes. In a largely unpopulated music industry or sector of that industry, fostering entry into the market will lead to large increases in the amount of music produced because producers will be initially tempted by the potentially supracompetitive profits available in a monopolistic environment. However, those

incentives are naturally limited by profit considerations. As more music is produced, albums become better substitutes for one another, creating a more competitive environment and reducing monopolistic pricing. The more legal policy facilitates the creation of close substitutes along the range of consumer tastes, the more the creation of new works leads to an increasingly competitive environment.

Furthermore, the stimulation of entry into the market and the facilitation of the creation of close substitutes increases access by lowering deadweight loss. Increased entry of close substitutes creates cross elasticity that flattens the demand curve for each particular work produced. The flattened demand curve decreases the price and quantity differential from surplus, and thus access, maximizing optimum.

Thus, the analysis demonstrates that the traditional notion of the access/incentive tradeoff is somewhat of a false construct. Though there may be a certain amount of deadweight loss, and thus reduced access, endemic to the market for music, those losses are not as great as many suggest and can be greatly minimized by fostering competition through entry into the market. Moreover, the access/incentive tradeoffs are essentially no different than in the case of other tangible goods. As in perfectly competitive markets, the cost of production is the primary limiting factor to access to music.

Having engaged in a thorough analysis of the general market dynamics, it seems prudent to take stock of the effects the copyright exceptions and structural limitations will have on those dynamics. As the analysis in section B suggests, the exceptions may be viewed as attempts to cure various market failures. The problem with justifying each exception on a case by case basis is that the justifications don't consider how the various exceptions, let alone copyright's structural limitations, interact with one another. It would seem more plausible that a combination of exceptions and limitations lend themselves to curing (or causing) various market problems. Thus, it seems prudent to lay out a number of possible market failures and then consider the effects of the exceptions and structural limitations on those market failures. To that end, let us recall and recast some of the market failures described previously, as well as introduce some additional ones.

1. Transactional costs: Transactional costs might be so high as to deter or prohibit valuable uses of a work. Implicit in that claim is the notion that the value of the use to the consumer would outweigh the costs, except that cost of contacting, negotiating, or otherwise engaging in the process of exchange would be cost prohibitive. As was previously explained, fair use provides a partial fix to this problem. Allowing consumers to engage in some free (fair) uses at the producer's expense increases total surplus because the value exceeds the costs. Additionally, structural limitations on the enforceability would seem to provide a solution for similar reasons.

2. Extension of monopoly pricing: As discussed extensively, monopoly pricing creates deadweight loss. The first sale doctrine ensures record companies cannot extend that monopoly by continuing to control price after an album is first sold. As a result, competition is bolstered in the retail market. The rise of a secondary market for used albums further bolsters competition and provides value to those who would otherwise forego consuming music. Furthermore, exceptions for radio transmissions and non-subscription webcasting provide at least nominal access to those who would otherwise forego music consumption. The result is an increase in total surplus and thus access to works.

3. Imperfect consumer information: Without accurate information about the price and characteristics of a good, consumers may make errors in their purchases. On one hand, consumers may mistakenly buy a good for which they have overestimated its value. On the other hand, consumers may forego buying a good for which they have underestimated its value. In either case, total surplus diverges from the ideal. Fair use provides a partial remedy to this issue. In the case of music, the effects of imperfect consumer information are potentially large because access to the work itself is often the only means to accurately assess value. Nevertheless, there are many ways of amassing information about music. The fair use exceptions for commentary, criticism, parody, research, and scholarship all result in an increase in information and ensures greater appreciation of music. Similarly, the exceptions for radio broadcasts, non-subscription webcasters, libraries and archives provide information by allowing limited, partial or time-constrained access to works. Additionally, the first sale doctrine provides a partial fix by ensuring consumers can lend purchased works to others (e.g., lending the CD to a friend so he can listen to it and determine whether it is worth buying).

4. Positive externalities from consumption: Consumption of artistic works often provides benefits to third parties. Unfortunately, the private market has no way of assessing the value to those parties. As a result, underproduction occurs where the marginal benefit to third parties outweighs the marginal cost to the producer. The result is an inefficient allocation of resources and decline in surplus from the potential maximum. This market failure differs from the failure resulting from imperfect consumer information because the value inheres in the actual consumption of the work itself, not the increase in information about the work. Typically, the solution to this market problem is to subsidize the use in some form. Fair use subsidizes consumption that creates positive externalities by providing free (fair) uses, at the producer's expense, when the combined value to the consumer and third party exceeds the costs to the producer. The result ensures third parties benefit from consumption by a second party whose use is fair (e.g., parents enjoy the school's fair use of a CD as musical accompaniment to a school play). Additionally, the first sale doctrine provides a partial fix by ensuring consumers can lend or give purchased works to others (e.g., lending the CD to the community center for use in the community center dance performance).

5. Artistic/Informational barriers to entry: Artistic works, like scholarly works, are largely derivative. Access to and utilization, either consciously or subconsciously, of pre-existing works is a necessary precondition to the creation of many types of works. However, artists cannot know, a priori, which existing works will facilitate or best support the creative process.⁴⁸ Thus, they cannot accurately assign a value to access and utilization of existing works in terms of its effects on future creation. What differentiates this market failure from the imperfect consumer information market failure is that no amount of information about the work, short of access *and* utilization, will cure the problem. Thus, potential artists may create less valuable works or forego creation at all. The result is the underproduction of valuable works.

All of the exceptions to, as well as the structural limitations of, copyright law's grant of exclusive rights serve to subsidize in part the costs of creation. This subsidization lowers the artistic/informational barriers to entry. Fair use does this by promoting access and derivative work creation for purposes of commentary, criticism, parody, research, and scholarship. The first sale doctrine subsidizes the cost of creation by lowering access and educational costs through the promotion of competition in the market. The exceptions for radio transmissions, non-

⁴⁸ Julie Cohen, Copyright and the Perfect Curve, 53 Vand. L. Rev. 1799, 1809 (2000).

subscription webcast services, reproduction for personal use, and library and archival use foster access and utilization of existing works. Finally, structural limitations to copyright enforcement further increase access to and utilization of works.

Now that the primary market failures have been identified, we have created a relatively complete methodology on which to assess the economic impact of technology on the market for music. As a general principle, technology that stimulates entry into the market and fosters the creation of close substitute works will increase total surplus. Additionally, technology that lowers transactional costs, impedes extension of monopoly power, perfects consumer information, subsidizes uses with positive externalities, and lowers the artistic barriers to entry will aid in correcting for market failures, thus increasing total surplus as well.

V. Economic Effects of DRM

With the preceding sections in mind, let me make a bit of a prediction here. The use of digital technologies will likely become a dominant, if not the predominant, method for the production and distribution of music within the next decade. Whether digital technologies will be employed in a music industry that creates substantial or little social welfare is quite a different matter. Admittedly, the music industry is still searching for a business model that will incorporate the advantages of digital technology while minimizing the costs of the disadvantages. Indeed, DRM technology is one attempt at developing a technological solution to those disadvantages. But regardless of whether DRM is ultimately accepted from a business and legal standpoint, history suggests that technology will ultimately win regardless of whether a particular industry embraces it or not. The question is not whether digital technology will be used, but rather to what degree, in what form, and toward what greater social end.

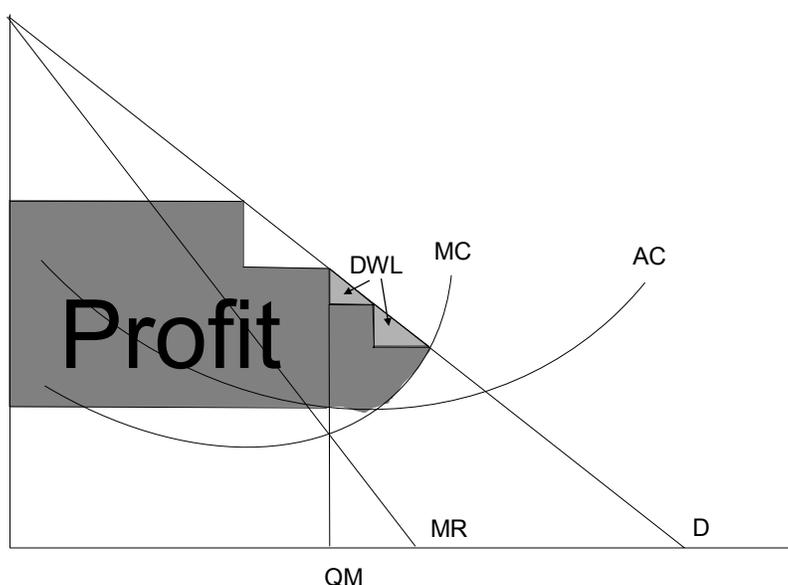
Thus, the remainder of this paper will take as its starting point the economic state of the music industry prior to the development of the digital technologies outlined above (with the exception of the CD itself), and will measure the relative increase in welfare based on how well the DRM solution utilizes digital technology's advantages and minimizes the costs from its disadvantages. Determining how successful the DRM solution is, as an empirical matter, relative to other possible solutions is beyond the scope of this paper. As such, any judgment about DRM will be rather subjective. Nevertheless, the remainder of this paper will attempt draw conclusions about how well the DRM solution incorporates the promise of digital technology in general, suggesting whether the music industry and lawmakers should search elsewhere for solutions to the problems posed by piracy and file sharing.

Utilization of DRM technology would foster price discrimination. The reason has to do with micropayments. Recall that micropayments are small, automatic charges to customers for each use of a work that are the likely result of implementing DRM technology. As a result of such technology, producers would have the ability to provide multiple pricing schemes for different types of uses and forms for the same product. For example, a typical CD, which usually contains an entire album and includes packaging and a tangible good that is conveniently ported from one place to another, is priced much higher than a downloaded file from iTunes, which contains only a single song without packaging in a non-tangible form that is less conveniently ported (in a physical sense) from one place to another. One can imagine that producers will eventually utilize DRM to charge different prices depending upon how many times the user wishes to listen to a song or the complete album, how many personal copies the user wishes to make, how often the user wishes to port the music from one device to another, and whether the user wishes to utilize

or manipulate the work in the creative process (e.g., sampling). The ability to divide up the potential market in this manner is often referred to as second-order price discrimination.⁴⁹

Figure 6 is the graphical representation of the effects of price discrimination on the market for an album, again from the point of view of our intrepid record company. Notice that the company can now offer multiple prices along the demand curve. This has two important consequences. First, it increases total surplus. The reason is that the fragmentation of the market through the pricing scheme allows the company to charge consumers prices for certain types of uses below the single monopolistic price. In essence, the album priced below the single monopolistic price is merely a cut down version of the original, perhaps a single song only or encoded in a form that allows only a certain number of listens or limits its portability. The result is increased access to the work (albeit in a stripped-down version) and a decrease in deadweight loss.

Figure 6



The second consequence of price discrimination is that the producer appropriates more of the surplus in the form of profits. Notice that the shaded area representing profits under the demand curve is significantly larger in proportion to the surplus than in the previous figures. This is the result of providing different prices for various stripped-down versions (e.g., single song downloads), as well as possibly beefed-up versions (e.g., special edition CDs with previously unreleased songs).

⁴⁹ Contrarians will argue that many of the examples described are not, in fact, examples of second-order, or any order, price discrimination, but rather product differentiation. Second-order price discrimination, also known as quantity discrimination, is when the firm charges different prices for the same good based on the amount the customer buys. Under second-order price discrimination, all customers face the same price schedule. In reality, many of the examples presented are different versions of the good, arguably even different products entirely, and thus may not be reducible to a common unit. Nevertheless, intuition suggests that the effects of dividing up the market in the manner represented in the examples produces the same general results as pure second-order price discrimination.

With those consequences in mind, DRM would appear generally to promote social welfare by stimulating entry into the market. Potential producers would enter the market once they recognized the potential for increased appropriation of surplus. Increased entry into the market would increase the number, and by necessity the diversity, of albums. As firms produced more albums, the cross-elasticity of demand and resultant competition would push prices downward and decrease deadweight loss. Moreover, consumer demand would be created and enhanced as the diversity of albums better catered to consumer tastes. This would increase total surplus significantly.

It should be noted that the magnitude of the increase in total surplus is disproportionately dependent on whether the stimulation of entry is large enough to facilitate the production of close substitutes. Where companies in particular sectors of the music industry, corresponding to the different genres of music, exhibit monopolistic or oligopolistic behavior, and where other untapped sectors exist, new entrants will opt merely to produce works of a different genre where they themselves can behave monopolistically. For example, suppose that there are only a few companies producing albums in only a few of the many potential genres. Specifically, suppose company A produces country albums and company B produces rock and pop albums. Company C, recognizing the existence of a potential market for folk music and the substantial profits reaped by companies A and B in their respective markets, would produce folk music albums. On one hand this result would increase total surplus because company C would tap into unmet consumer demand. However, price inefficiencies would remain largely intact because the albums produced by companies A, B, and C would be poor substitutes for one another. Thus, the disproportionately high surplus gains that are attributable to cross elasticity and the resultant competition, as well as creation and enhancement of consumer demand, would never materialize.

In order for DRM to significantly increase surplus, the stimulation of entry caused by price discrimination would have to outweigh the barriers of entry into existing markets for existing genres of music. Where entry costs are potentially very high relative to potential profits or where dominant firms can engage in predatory pricing, DRM may provide only minor increases in surplus through the partial elimination of deadweight loss.

Determining entry costs for the music industry is not terribly difficult. Part IV described the nature of both the fixed and variable costs for an album. Fixed costs associated with recording, marketing, designing, and administering an album have traditionally been quite large. However, as the section on digital technology described, costs for recording equipment and related software have dropped significantly. Furthermore, digital file formats and their related devices, as well as the Internet provide cheap alternatives to traditional reproduction and distribution. All of these advances provide the potential for increased cost savings, and thus the potential to lower the costs of entry into the music industry.

Up to this point, we have assumed a cost-free implementation of DRM. In reality, however, DRM implementation is not cost-free. Technological access controls must be continually updated. The costs associated with continual update could potentially create significant additional fixed and variable costs. From a theoretical standpoint, it is impossible to determine whether increased costs from DRM implementation will significantly outweigh the savings from utilizing the other digital technologies described above. However, it is important to note that the costs of DRM implementation would at least partially negate, if not overshadow, the savings from utilizing other digital technologies.

In theory at least, it is possible that DRM implementation will increase overall costs for production of an album. Herein lay the rub. If the loss in surplus due to increased variable costs outweighs the gain in surplus from the elimination of deadweight loss through price discrimination, total surplus might in fact decline.

Unfortunately, a decline in total surplus due to the costs of DRM implementation would not necessarily lead the music industry to forego the use of the DRM solution. This is due to the fact that price discrimination increases the ability of a producer to appropriate a larger proportion of the surplus created. Thus, it is possible that while total surplus might decline, a producer's profits might increase. An illustration is probably in order. Suppose that the total surplus from the production and consumption of an album without DRM technology is 100x and that the record company appropriates 50%, or 50x, of that surplus (profit). Now suppose that a record company utilizes DRM technology. Suppose further that the costs associated with DRM implementation outweigh the gains from the elimination of deadweight loss, resulting in total surplus of 90x. Finally, suppose that the record company can appropriate 60% of the surplus due to price discrimination facilitated by the DRM technology. The result is that the record company will appropriate 54x of the surplus (profit). Thus, it is possible that DRM implementation will decrease total surplus while increasing producer profits. Moreover, the aggregate effect of the decrease in total surplus across the entire market might be quite large.

This scenario should be cause for some concern. On one hand, increased profits might still stimulate entry into the market with a resultant increase in total surplus. On the other hand, the costs of DRM implementation threaten to negate that increase in total surplus while substantially raising barriers to entry. Referring back to the previous example, if the 50% appropriation level represents the market equilibrium, the 8% increase in profit garnered by the producer may not serve as enough of an inducement for other producers to enter the market, particularly if the costs of DRM implementation dramatically increase the overall cost of production. Contrary to some predictions, DRM may further entrench existing producers.

Having considered the general market effects of DRM implementation, it seems appropriate now to evaluate its potential effects on the various market failures outlined in Part IV. As it relates to transactional costs, DRM would appear to have a beneficial effect. The reason is that the automation of payment that DRM technology facilitates would decrease transactional costs. Assuming, as we did in Part II, that recognition of all or most of the potential uses could be coded onto the rule set that controls access to a digital file, the costs related to the process of exchange would radically decrease.

The potential for DRM to facilitate the extension of producer monopoly pricing power into the retail markets depends on the success of DRM in promoting competition in the music industry. The more DRM stimulates entry into the market and fosters the creation of close substitute works, the more DRM will lead to competition. The greater the resultant competition and corresponding deadweight loss, the less important producer control in the retail market will become. As the economic analysis suggests, it is uncertain as a theoretical matter whether DRM implementation will necessarily promote competition. To the contrary, it is quite possible that DRM will only serve to entrench existing producers.

If DRM does not facilitate competition, existing producers might extend their monopoly pricing power into the retail market and effectively shut down any secondary market for music. Nevertheless, this seems an unworthy cause for great concern for two reasons. First, even if

producers absolutely controlled the prices retailers charged, or took over retail sales themselves, the pricing strategies from the retail perspective would remain the same. Second, price discrimination will provide increasing access to customers who value works the least. It is in the interest even of monopolist to offer increasingly smaller units or cut-down versions to the lowest valuing customers in order to appropriate the largest amount of surplus. The iTunes DRM implementation provides a concrete example of why this outcome is not problematic. iTunes currently charges about \$1 to download a single song file. Other services are offering similar prices. At these prices, there appears to be little cause for concern. Moreover, it is hard to imagine a burgeoning secondary market in single song files that retail at \$1.

DRM will potentially diminish the capacity for the copyright exceptions to provide information to consumers. On one hand, DRM implementation would have no specific negative effect on radio broadcasters or non-subscription webcasters, as producers have a vested interest in designing DRM so that these services succeed. The reason is that radio and webcasting provide free promotion of the producers' albums. Similar arguments apply to libraries and archives, as well as the fair use exception for commentary, criticism, research, and scholarship. Moreover, DRM could not effectively control the ability of particular people to purchase the work and report back to the populace about the work's characteristics. On the other hand, consider the act of lending music to a friend so that he can decide whether to buy it. Lending arguably has a tremendous impact on consumer information in the aggregate. Assuming the technology could facilitate the lending of a song file (perhaps by deleting itself on the owner's computer once the file was transferred, and vice versa when the recipient was finished), it would nevertheless charge for the use. The reason is that there is no way for technology to determine the intended purpose of a transaction. It can only distinguish between each type of use. In order for the technology to be effective in the first place, it must charge for each use, at least the type of use engendered here. But in this case, and in other cases we will examine in a moment, the purpose is quite different than typical consumption. If the information and enjoyment obtained is not greater than price charged, then the transaction decreases total surplus. In each individual case, the decrease in surplus might be quite small. However, in the aggregate, the decrease might be quite large.

The code-plus method of DRM implementation does not provide a ready fix. Remember that in cases where a particular use is prohibited or the consumer wishes the charge to be waived, the code-plus method allows a user to apply for a digital access key and have the case judged by a human decision-maker. Unfortunately, in most instances the transaction costs of such a process would likely be significantly greater than the value of the information obtained from waiving the charges. Moreover, the incurrence of transaction costs would lower the level of surplus below that which existed before the implementation of DRM.

Additionally, DRM will have a negative impact on positive externalities from consumption. Recall that that this is when the use by one person causes beneficial effects on others, increasing total surplus. In certain cases, fair use provides a means for wholly subsidizing such use at the producer's expense because the use increases total surplus. The problem, again, is that there is no way for technology to determine the intended purpose of a transaction. Nor can it assess the effect of on third parties. Thus, it will charge for the transaction. That charge will prohibit or deter certain uses, for example, giving an album in a digital format to the school to provide accompaniment for the school play. Again, each loss of surplus might be quite small, but in the aggregate might be quite large. Moreover, any attempt to utilize the code-plus methodology will merely introduce transactional costs.

Finally, DRM implementation may lead to the underproduction of musical works. Recall that music creation is a derivative process and thus relies on access to and utilization of existing works. The perfect example of this phenomena is the now ubiquitous practice of sampling (another example is the creation of the mix tape or mix CD). By utilizing and manipulating snippets of other works and layering them along with the recordings of their own performances, artists can create new works. As was discussed in Part III, under the current legal regime, artists must pay the copyright holder before such a product is brought to market because such works qualify as derivative works. Importantly, artists generally are not currently charged for the utilization of the sample in the creation process, but rather only when the resulting work is brought to market. Unfortunately, DRM implementation would likely change that scenario. The reason is that producers will choose a level of protection and a pricing scheme that protects their investment in, and creates returns from, their existing works. Thus, once a producer creates a work, it has every incentive to charge for as many uses as possible. Interestingly, that incentive would likely lead to producers charging artists at the point of creation, not at commercialization of the existing work (although an additional charge might occur there as well). Even if the price of sampling was relatively small, say \$.01, certain artists might forego creating music because they could not, a priori, determine the value of any particular sample in the recording process. Thus, DRM implementation has potentially tremendous negative repercussions for the music industry.

Even if producers in the music industry decide to waive such fees amongst themselves (or amongst the artists they employ), they are still likely, for the previously stated reasons, to charge hobbyists or other artists not recording in their studios. This is of considerable importance for two reasons. First, potential artists must practice sampling to develop skills in the art of sampling. However, charging for the use of samples would disproportionately affect those learning to create music. Thus, charging for the use of samples would increase barriers to entry. Second, while artists that eventually bring their works to market might recoup the costs of sampling charges, the artists, hobbyists, or experimental-minded consumers who do not commercialize their works cannot. It is this non-commercial sector of the artistic community that is not valued on any of the economic graphs created thus far. However, given the abundance of freely available art (consider the ubiquitous mix tape or mix CD which, yes, is itself a work of art), there is reason to believe that the lost surplus caused by artists foregoing artistic creation will be quite large.

VI. Conclusion

Having examined the theoretical economic effects of DRM implementation on the music industry, we are left with a rather uncertain future. DRM may indeed tend to increase competition and diversity in the music industry. However, it is also quite possible that DRM implementation will create losses in total surplus in the music industry. Empirical analysis is necessary to study whether the DRM solution can be implemented cheaply enough to lead to greater competition. Furthermore, the analysis suggests DRM will impede some of copyright law's attempts at curing various market failures. Legislation might cure part of that problem, for example by mandating the free use of samples within DAWs. However, the analysis also suggests that exacerbation of some market failures may be endemic to the DRM solution.

Perhaps it is time to search for another solution.