Intellectual Property Strategy and the Long Tail: Evidence from the Recorded Music Industry

Laurina Zhang*

October 27, 2014

Abstract

Digitization has impacted firm profitability in many media industries by lowering the cost of copying and sharing creative works. I examine the impact of digital rights management (DRM) - a prevalent strategy used by firms in media industries to address piracy concerns - on music sales. I exploit a natural experiment, where different labels remove DRM from their entire catalogue of music at different times, to examine whether relaxing an album's sharing restrictions increases sales. Using a large sample of albums from all four major record labels, I find that removing DRM increases digital music sales by 10% but relaxing sharing restrictions does not impact all albums equally. It increases the sales of lower-selling albums (i.e., the "long tail") significantly (30%) but does not benefit top-selling albums. These results suggest that the optimal strength of copyright depends on the distribution of products in firms' portfolio.

*Ivey Business School, Western University, 1255 Western Road, London ON, N6G 0N1, Canada, lzhang@ivey.uwo.ca
Acknowledgments: I am grateful to my PhD advisors for their insights and guidance over the years: Ajay Agrawal (chair), Joshua Gans, Avi Goldfarb, and Joanne Oxley. I also thank Brett Danaher, April Franco, Alberto Galasso, Shane Greenstein, Doug Hyatt, Ariel Katz, Nicola Lacetera, Matt Mitchell, Carl Shapiro, and participants of NBER Summer Institute's Digitization meeting and Searle Roundtable on the Law & Economics of Digital Markets for helpful suggestions. I gratefully acknowledge funding support from the NBER Economics of Digitization and Copyright Initiative, the Martin Prosperity Institute, and the Centre for Innovation and Entrepreneurship.

1 Introduction

Digitization has led to Schumpeterian creative destruction (Schumpeter, 1934) in many media industries (e.g., music, books, movies) by significantly lowering the cost of copying and sharing creative works. At the same time, new production and search technologies have changed the quantity and variety of products available for consumption, often to the benefit of niche products at the tail of the sales distribution (Anderson, 2004, 2006; Brynjolfsson, Hu, and Simester, 2011). Consequently, digitization significantly impacted firm appropriability in a variety of settings, in many cases shifting surplus from producers to consumers. A central issue facing firms in this setting is how to manage their intellectual property (IP) strategy for digital products such that incentives to facilitate product discovery, given the increasing quantity and variety of products available, are balanced with incentives to restrict illegal file sharing (i.e., piracy). While the government traditionally defines and enforces copyright laws, more firms increasingly rely on technology to assert property rights and combat piracy.

The recorded music industry is perhaps the most visible example of an industry impacted by such technological changes. The emergence of file sharing technology has provided consumers with the ability to cheaply reproduce digital music files and disseminate them across the globe using peer-to-peer networks. The Recording Industry Association of America (RIAA) estimates that music sales declined by half in ten years - from $14.6 billion in 1999 to $7.7 billion in 2009 (RIAA, 2010) - and attributes the decline to piracy.

While the recording industry has experimented with varying legal responses to these changes over the past decade, from suing file sharing services and individual users to lobbying for stronger copyright enforcement, few of these actions have led to permanent changes to the legislation.[1] Whereas these past efforts have relied on IP policy, record companies

[1]One notable exception is the *MGM Studios v. Grokster* Supreme Court decision in 2005 where Grokster was forced to shut down their file sharing site and pay $50 million to the recording industries (http://archive.today/3Gea [accessed September 2, 2013]). The resulting decision set the precedent such that producers of technology who promote the ease of copyright infringement can be be held liable, essentially allowing the RIAA to go beyond merely suing individuals who share files illegally to suing the companies whose software enables the sharing.

in recent years have instead turned to encryption-based digital rights management (DRM) technologies to assert property rights (Lichtman, 2006). DRM technologies allow publishers and copyright holders to exert control over how consumers use digital content by making it difficult, if not impossible, to reproduce and distribute copies of legally purchased digital music. Thus, DRM technology is an IP strategy implemented by copyright holders to make digital content excludable through a combination of technical restrictions and legal enforcement. While DRM raises the cost of piracy, its sharing restrictions may also raise search costs and hinder product discovery. A central issue underlying copyright policy debates in light of new technology is how to balance the incentives for diffusion of creative products with the incentives for legal purchases.

In this paper, I examine the impact of firms' IP strategy (DRM) across the distribution of music sales (i.e., mainstream vs. niche music). The extent to which firms' IP strategy can differentially impact the sales of different types of products in their portfolio remains relatively unexplored and has strong firm-level implications. Specifically, the extent to which DRM impacts music sales is likely to be different for music from well-known and lesser-known artists. Given that sharing allows consumers to gain information about the product fit before purchase (Chellappa and Shivendu, 2005; Peitz and Waelbroeck, 2006) and prioritize information in settings where consumer attention is scarce (Gans, 2012), sharing restrictions are more likely to hinder the discovery of lesser-known music. Conversely, sharing restrictions are unlikely to increase consumers' purchase of music from well-known artists because they have likely experienced and determined their preference for the music beforehand (e.g., through radio). Thus, DRM's countervailing effects that stem from piracy and product discovery can have differential impacts on the sales of music at different parts of the sales distribution.

The empirical context of the paper is the four major record companies - EMI, Sony, Universal, and Warner - removing DRM from their entire catalogues of music at different times. Specifically, EMI drops DRM from their catalogue in 2007, while the remaining major labels do not completely remove DRM until 2009. I construct a large sample of albums from all four major record companies (some with DRM removed, and some not) covering multiple genres and different parts of the sales distribution for the years 1992-2011. The sample comprises of 5,864 albums from 634 artists and is, to my knowledge, the longest and broadest panel constructed to describe music sales. The data includes album-month level data on the number of albums sold through physical (e.g., WalMart) and online (e.g., iTunes) channels.

The main empirical challenge associated with isolating the impact of DRM on sales is that there may be unobservable album-level heterogeneity and unobserved EMI-specific factors that are correlated with DRM removal. Three parts of my empirical strategy help alleviate these challenges. First, I implement a difference-in-differences estimation where I compare the sales of similar albums with (non-EMI) and without DRM (EMI) to identify how the level and distribution of music sales change after the removal of DRM. Second, institutional details suggest that EMI's decision to remove DRM was relatively unanticipated and EMI did not make the decision to drop DRM in anticipation of disproportionate changes in sales to any part of their catalogue. I present time-varying estimates to show that there is no evidence of pre-trends in the sales of EMI albums before DRM removal. Third, my estimations on the impact of DRM removal across different parts of the sales distribution relies only on variation within EMI's catalogue of albums (rather than variation across labels). This mitigates concerns that unobserved EMI-specific factors are driving the difference between EMI's top-selling and lower-selling albums. In other words, the core identifying assumption is that for each individual album released before 2007, EMI's decision to drop DRM is exogenous. Thus, the focus of this paper, and where the exogeneity of the DRM "experiment" is the strongest, is on the impact of DRM removal on changes to different parts of the sales distribution.

My estimates suggest that removing DRM increased digital music sales by 10%. Importantly, the impact of DRM removal is not uniform across the sales distribution. I find that relaxing sharing restrictions disproportionately increases sales of albums in the long tail (i.e., lower-selling) albums significantly (30%) but does not benefit top-selling albums.

While a potential concern is that increases in the sale of long tail albums may be due to the higher value of DRM-free music, I do not find evidence that the increase is attributed to changes in value or piracy. I provide support that DRM removal facilitates the discovery of lesser-known music by exploiting variation in artists that have released albums under multiple major labels prior to DRM removal (e.g., Al Green who has released albums under both EMI and Universal). If the discovery mechanism holds, then dropping DRM on the artist's EMI albums should also lead to an increase in the sale of its albums released by other labels, even though the value of the artist's non-EMI albums has not increased. I show that DRM removal increases the sale of EMI artists' non-EMI albums in the long tail but does not benefit its top-selling non-EMI albums.

In addition, I find that the increase in sales of lower-selling albums is not just driven by the sale of older albums. I find that tail albums of newer vintages experience a comparable increase in sales relative to tail albums of older vintages. Furthermore, lower-selling albums of less pirated genres (e.g., Jazz and Classical) disproportionately benefit from relaxed sharing restrictions compared to actively pirated genres (e.g., Hip Hop and R&B). I interpret increased sharing as lowering search costs, and as such, my results are consistent with theory that shows lowering search costs can facilitate discovery of niche products in the long tail.

This study offers three main contributions. First, departing from prior studies that estimate the magnitude of sales displacement from piracy (Rob and Waldfogel, 2006; Zentner, 2006; Oberholzer-Gee and Strumpf, 2007) and the impact of changes in copyright law (Png and Wang, 2006; Danaher, Smith, Telang, and Chen, 2013), I provide the first empirical evidence (to my knowledge) on the relationship between digital content sales and IP as a result of firms' strategic decisions. I show that firms' IP strategy differentially impacts the sales of mainstream vs. niche products, suggesting that the optimal strategy for IP in creative industries depends on the distribution of products in firms' portfolio. Given the unpredictability of product appeal at the long tail (Aguiar and Waldfogel, 2014), this finding is consistent with the view that copyright institutions governing online markets that facilitate discovery and diffusion of digital goods is potentially surplus-enhancing. Second, given that DRM is currently implemented and debated in other settings (e.g., books, movies, video games), my study helps to inform other settings that are undergoing similar transitions in market structure and competitive behavior. More broadly, I contribute to the vibrant literature on IP by considering its impact on sales, thus complementing existing research that has focused on its impact on knowledge flows[2] (e.g., Agrawal and Henderson, 2002; Jaffe, Trajtenberg, and Henderson, 1993), cumulative innovation (Murray and Stern, 2007; Williams, 2013; Galasso and Schankerman, 2013), and price (Li, Macgarvie, and Moser, 2013).

The paper proceeds as follows. The next section provides an overview of the recorded music industry and the relevant literature. Section 3 provides details on the DRM technology and the DRM "experiment." Sections 4 and 5 describe the data and empirical approach. Section 6 presents the results, and Section 7 presents concluding remarks.

2 Setting & Related Literature

Creative products, such as music, movies, and books, have high fixed costs and low marginal costs of production. Private firms have been able to profitably bring these products to market because they are excludable through a combination of technology and a complementary legal framework provided by copyright law. Given that competitive markets may under-incentivize innovation because of the public-good nature of ideas (Nelson, 1959; Arrow, 1962), IP rights, such as patents and copyrights, aim to incentivize innovation by allowing firms to capture a higher share of the social returns to their research investments.[3] While the local monopoly granted to creators gives rise to that monopoly's usual harm to consumers, this harm is thought to be offset by copyright's incentive effects on the creation of new works.

[2]See Moser (2013) for a review.

[3]Patents and copyrights differ in several important ways. First, patents in general offer a broader scope of protection compared to copyrights. Patents protect the physical process and the invention and give the inventor the right to exclude others from making, using, offering for sale, or selling the invention, while copyrights protect the expression of unique ideas, including literary, dramatic, musical, and artistic. However, copyright does not protect the underlying subject matter of the expression. Second, copyrights offer longer term of protection compared to patents. In the U.S., patents protect the invention for a minimum of twenty years, while copyright provides protection for the life of the author plus seventy years. Interestingly, there is a negative externality associated with patents and copyrights both in terms of their impact on cumulative innovation. McLeod and DiCola (2011) note that copyrights may restrict musicians building on another's prior works and there is a cost associated with figuring out the samples that require licenses under copyright law and then negotiating the license fee with each copyright owner.

Indeed, the International Federation of the Phonographic Industry (IFPI) states: *"music is an investment-intensive business ... Very few sectors have a comparable proportion of sales to R&D investment to the music industry."* RIAA also states: *"all these anti-piracy efforts are to protect the ability of the recording industry to invest in new bands and new music."*[4]

The advent of digital media and analog/digital conversion technology has materially lowered the costs of copying and sharing in the music industry and, as a result, has vastly raised concerns about effective copyright protection. The advent of personal computers has made it convenient for consumers to convert media in physical form (i.e., CDs) into a digital form through ripping, and most notably, peer-to-peer networks have made sharing and copying music, a once cumbersome and time-consuming process, essentially costless.[5] Most observers agree that the technological change since the late 1990s has sharply reduced effective copyright protection for music. The music industry has been unequivocal in blaming file sharing for the decline in sales and argues that piracy has serious consequences for whether new works will be brought to market. The recording industry has sued numerous file sharing services as well as thousands of individuals who share files.[6]

A growing body of literature has focused on the extent to which file sharing has displaced album sales. Theoretical literature has shown that piracy is not definitively bad for firms. For instance, piracy may be beneficial for a new product if the firm needs to establish an initial user base to speed up diffusion (Prasad and Mahajan, 2003), and piracy can act as a free "sample," increasing product awareness (Peitz and Waelbroeck, 2006; Gopal, Bhattacharjee, and Sanders, 2006). The empirical literature is mixed, with most studies finding some displacement in album sales (Blackburn, 2004; Liebowitz, 2006; Rob and Waldfogel, 2006; Zentner, 2006). Several papers using direct measures of file sharing do not find evidence that file sharing significantly affects sales (Oberholzer-Gee and Strumpf, 2007; Smith and Telang, 2009). There is also little evidence of an aggregate decline in either the quantity or quality

[4]See http://www.ifpi.org/content/section_news/investing_in_music.html [accessed September 10, 2012].

[5]It is interesting to note that copying technologies have always been disruptive to market structure and competitive dynamics of a "Schumpeterian" manner. Player piano rolls early in the 20th century, audio tape recording, and video tape recording had always been objected to by copyright holders and content producers. See Scherer (2003).

[6]http://www.wired.com/2010/05/riaa-bump/ [accessed June 2013]

of recorded music produced (Waldfogel, 2011, 2012).

However, few of the previous studies explicitly take into account the emergence of legal online markets, which have arguably shifted substitution between file sharing and legal products and also have implications for the distribution of goods consumed. The emergence of legal online platforms, such as the iTunes Music Store, provides consumers with different options for music consumption. In addition to having a legal digital alternative, which may substitute for offline purchases (Goolsbee, 2001; Ellison and Ellison, 2006; Prince, 2007; Forman, Ghose, and Goldfarb, 2009), consumers also have more options to buy less music since albums are "unbundled" online (Bakos and Brynjolfsson, 1999). Moreover, consumers' modes of music discovery, previously dominated by local radio,[7] now include global streaming sites, social sharing, and recommendation engines.[8] These additional channels are able to help consumers locate, evaluate, and purchase a far wider variety of products than they can via traditional brick and mortar channels.

Digitization also has impacted the distribution of products available for consumption. The long tail argument suggests a sharp increase in the variety of products offered through online channels, which fuels a shift in consumption away from hits to a much larger number of lower-selling niche products. For example, whereas Wal-Mart may only carry the top 3000 music albums that have the broadest mainstream appeal due to limited shelf space and local demand, online retailers, like iTunes, Amazon, and Rhapsody, can profitably carry niche albums with limited appeal because the cost of stocking an additional album on the Internet is virtually zero and online retailers can aggregate demand by finding audiences across the globe. This argument suggests that niche content, such as older catalogues, music from indie artists, and remote genres, are able to find an audience and earn similar margins to a "hit" album (Anderson, 2004, 2006). The long tail effect of digitization has been documented in book sales (Brynjolfsson, Hu, and Smith, 2006), home video sales (Elberse and Oberholzer-Gee, 2007), and music consumption (Bhattacharjee, Gopal, Lertwachara, Mars-

[7]Hendricks and Sorensen (2009) quantify the extent to which albums "lost" sales because consumers may not have known about them in the period before the emergence of online markets.

[8]For example, Dewan and Ramaprasad (2012) find that online social media positively affects music sampling, particularly at the tail of the sales distribution.

den, and Telang, 2007; Dewan and Ramaprasad, 2012). Missing from the debate, however, is the link between the sales distribution of digital content and firms' private response to changes in the IP environment.

3 The DRM "Experiment"

DRM systems are access-control technologies used by hardware manufacturers, publishers, copyright holders, and individuals to exert control over the use of digital content and devices after sale. For music publishers, DRM offers the technical means to exert control over the use and distribution of digital music by making it difficult, if not impossible, to reproduce and distribute copies of legally purchased digital music. As the cost of copying digital content becomes lower, content vendors have started to use technical protection rather than simply relying on traditional legal frameworks on copyright as a means to curb piracy. In 1998, the Digital Millennium Copyright Act (DMCA) was passed in the United States to impose criminal penalties on those who make available technologies whose primary purpose and function is to circumvent content protection technologies. In other words, DRM is an IP strategy implemented by copyright holders to make digital content excludable through a combination of technical restrictions and legal enforcement. Currently, most firms in creative industries use DRM to address piracy concerns.[9]

The four major record companies[10] - EMI, Sony, Universal and Warner - which control the distribution of over 80% of the music market (International Federation of Phonographic Industry, 2005), first required that DRM systems be implemented in conjunction with the

[9]For example, DRM on e-books is currently in the process of being removed by book publishers, while the film and video game industries want to add more stringent DRM to their catalogue.
http://kevinomclaughlin.com/2012/04/25/publishers-begin-removing-drm-from-ebooks/
http://www.forbes.com/sites/insertcoin/2012/02/29/hollywood-wants-to-use-gamings-drm-to-protect-hd-movies/
http://www.yalelawtech.org/ip-in-the-digital-age/the-video-game-industry-and-drm-time-for-a-change/
[accessed August 2012].

[10]Note that record companies may be small, localized, and "independent" (indie), or they may be part of a large international media group, or somewhere in between. As of 2011, only four record companies can be referred to as "major." EMI was acquired by Universal in September 2012. However, my sample ends in June 2012, so the concern that the acquisition had an effect on my treatment is mitigated.

emergence of online music markets to protect their music from being illegally copied. It was not until viable DRM-protected music services, such as Apple's iTunes Music Store, were launched that consumers had a brand-name outlet to purchase music for download from the major music record labels. However, use is limited, as much as possible, to the individuals who make a purchase. For example, if you purchased a song on iTunes in 2003, you could have transfered it to any iPod in your household, but it was harder to transfer it between computers and certainly beyond.

While content providers, such as record companies, claim that DRM is necessary to fight online copyright infringement and prevent IP from being copied freely, those opposed to DRM argue that its restrictions do little to prevent copyright infringement and make it costly for legitimate consumers to use their legally purchased music. In particular, digital locks can restrict users from engaging in activities that are perfectly legal, such as making backup copies of a song, lending materials out to friends and family, accessing works in the public domain, or using copyrighted material for research and education under fair use laws. Sharing restrictions also can hinder the extent to which consumers can discover new music, since sharing technologies and recommendations are an efficient way for consumers to discover and purchase products that they otherwise would not have considered (Brynjolfsson, Hu, and Simester, 2011).

In addition, since DRM systems are usually proprietary to the service provider, music purchased from one vendor, such as Microsoft's Zune, are not playable on other devices, and content can become permanently inaccessible if the DRM scheme changes or if the service is discontinued. All of these restrictions impose costs on consumers (Sinha, Machado, and Sellman, 2010; Vernik, Purohit, and Desai, 2011). Thinking of the long term, some consumers might make investments to pirate music just to be free of the later hassles and restrictions caused by DRM.[11] At the extreme, opponents of DRM argue that it may stifle competition and decrease social welfare (Petrick, 2004). Thus, the effect of DRM removal on album sales

[11] For instance, consumers who purchased music may buy new laptops in the future and have to re-authorize their accounts. They also may acquire new family members to whom they cannot transfer songs. Analogously, consider a consumer who purchased a DVD but is subjected to an ad warning against piracy every time she views it (Gans, 2012).

and, in particular, on the sales distribution, is ambiguous and an empirical question.[12]

A major change occurs in April 2007, when EMI becomes the first major record company to remove DRM protection on its entire catalogue of music. In my interview with EMI's former Chief of Digital Operations, Barney Wragg, who was responsible for dropping DRM, EMI made this decision because they believed that *"DRM was making it cumbersome for consumers to use music the way they would like, and there was never a form of DRM that can protect against flagrant piracy."* Importantly, Wragg stated that EMI did not make this decision because of their catalogue composition. In other words, EMI did not make the decision to drop DRM in anticipation of disproportionate changes in sales to their long tail catalogue.

EMI's decision to remove DRM came as a surprise. In fact, when EMI made the announcement to drop DRM in April, many speculated that it was an April Fool's joke.[13] The decision was also controversial both within EMI and the industry. Wragg recalled: *"The other labels were surprised by this move. I was basically ostracized - I was not invited to anymore committee or board meetings within the industry."* This is largely because the major record companies have traditionally been staunch supporters of DRM technology. Other major record companies have openly critiqued the idea of removing DRM from their offerings, arguing that the technology will become increasingly important once digital sales eclipse CD sales. Edgar Bronfman Jr., chairman and CEO of Warner Music Group, famously argued: *"I don't agree that intellectual property should have no protection. We should all agree that intellectual property deserves some measure of protection."*[14] Even after EMI made the announcement to drop DRM, commentators did not expect the other record companies to jump on the bandwagon.[15]

[12] The theoretical literature on piracy and sharing suggests that while piracy can create an illegal source of competition and have a negative impact on rights-holders' profits, it may also increase the rights-holders' profits through sampling and network effects. Specifically, literature on DRM shows that whether being DRM-free is optimal depends on the level of copyright enforcement and the strategic interaction among producers of digital goods. Similarly, piracy is shown to hurt superstars, to the benefit of niche and young artists. See Belleflamme and Peitz (2011) for a review.

[13] http://www.guardian.co.uk/technology/blog/2007/apr/01/emiandapplea [accessed October 15, 2012].

[14] http://www.macworld.com/article/1055379/warner.html [accessed August 20, 2013].

[15] http://www.wired.com/entertainment/music/news/2007/04/emi_business0403 [accessed August 20, 2013].

EMI's decision to remove DRM means that any music purchased online that is owned by EMI can be copied and shared among friends and playable across different devices. This also means that putting a song up on a file sharing service and letting friends download it is now possible (though still illegal). Thus, DRM removal highlights the tension between increasing online consumption and potentially facilitating piracy. The other labels do not completely abandon DRM until April 2009, when Apple, which controls more than 80% of online music sales, negotiated deals with the remaining three major labels to have their content on Apple DRM-free. Thus, the removal of DRM has been enacted at different times across the major record companies.[16]

4 Empirical Strategy & Identification

In order to evaluate the effect of DRM removal on sales, I am faced with a fundamental inference problem. For a given album where DRM is removed ("treated" album), I cannot observe the counterfactual - the changes in sales if DRM is not removed on the album. Ideally, I would assign albums randomly across two groups and remove DRM on one group to disentangle the marginal effect of DRM removal on album sales. While I cannot replicate this ideal experimental design, I develop an empirical strategy that takes advantage of several features of my institutional setting to isolate the marginal impact of DRM removal on both the level and distribution of album sales.

The identification of the causal effect of DRM removal on digital sales would be difficult had all the major record companies lifted DRM at the same time. This is because any changes in sale may be attributed to pre-removal time trends and omitted variables such as unobserved album-level heterogeneity. However, because the major record companies removed DRM at different times, with EMI removing DRM in April 2007 and the remaining

[16]The three remaining labels released part of their catalogue DRM-free on Amazon's music store in the fall of 2007, but the coverage of albums is most complete on iTunes, and I do not expect to see significant changes in digital sales from the presence of Amazon. Empirically, I am not able to identify which albums were available DRM-free on Amazon, but I include month-year fixed effects in my main specification to control for the emergence of the Amazon music store. In the robustness section I also explore the possibility that the presence of DRM-free tracks on Amazon have contributed to changes in sales.

three major record companies removing DRM in April 2009, I can employ a difference-in-differences (DD) strategy, where I compare the sales of similar albums with and without DRM before and after DRM removal. My main estimating equation is:

$$log(OnlineAlbumSales{+}1)_{it} = \alpha{+}\beta_(EMI_i{\times}\ PostDRMRemoval_{it}){+}\sum_{s=1}^{6} w_s t^s{+}X_{it}{+}\delta_i{+}\mu_t{+}\epsilon_{it} \quad (1)$$

where I define $OnlineAlbumSales$ as the number of digital tracks sold by album i in month t divided by 10, following industry standards.[17] This is an appropriate measure of digital sales because the majority of sales of online music are from digital downloads. I log the dependent variable because album sales are skewed. Since the specification relates log of sales to the dummy variable $EMI_i \times DRMRemoval_{it}$, I compute the marginal effect as $e^{\beta} - 1$.[18] EMI_i is a dummy variable equal to 1 for albums released by EMI. $PostDRMRemoval_{it}$ is equal to 1 after EMI drops DRM in April 2007 for all albums to capture counterfactual changes in album sales had other labels dropped DRM at that time. Thus, β captures the marginal effect of DRM removal on album sales. Given that all major labels remove DRM by April 2009, my estimate is only identified until April 2009.[19] Therefore, I trim my sample at April 2009.

Making use of the fact that I observe online sales for 70 months, I can control for album (δ_i) and month-year fixed effects (μ_t). Album fixed effects control for all time-invariant differences between albums, such as genre and vintage. Month-year fixed effects control for changes over time that affect all albums similarly, such as economic downturns or the emergence of Napster in 1999. I also include a polynomial time trend of degree six (t^s), where t denotes time in months, to flexibly control for differences in sales due to album release

[17] http://www.billboard.biz/bbbiz/industry/retail/the-2011-music-sales-boost-by-the-numbers-1005339412.story [accessed March 2012].

[18] This follows because in the semilogarithmic model $lnY = \beta D$, where D is a dummy variable, $\frac{Y_1 - Y_0}{Y_0} = e^{\beta} - 1$ where Y_1 and Y_0 are the values of the dependent variable when D is equal to 1 and 0, respectively. Since the dependent variable is logged plus one, there is also a bias in interpreting the marginal effects. However, the bias understates the true effect and goes to 0 quickly (bias is - $\frac{\Delta Y}{(Y+1)Y}$)

[19] Estimation beyond April 2009 is only identified by variation in month-year fixed effects.

dates. Robust standard errors are clustered by album in order to reduce the potential for overstating statistical significance due to serial correlation within albums (Bertrand, Duflo, and Mullainathan, 2004).

In order to interpret my coefficients as an average treatment effect my identification strategy assumes that the timing of DRM removal is uncorrelated with factors that determine the outcomes of interest, conditional on the baseline controls. This assumption captures the fact that EMI's decision to remove DRM is a label-level decision that is not correlated with the sales of any particular album in the catalogue before DRM removal. I start by taking the identifying assumption as given and then check the validity of my assumption in the robustness section. While anecdotal evidence suggests EMI's decision to remove DRM comes as a surprise, there is still the concern that the timing of DRM removal is endogenous, such that the determinants of timing are correlated with factors that could affect the outcomes of interest (online sales) through channels other than DRM removal. If EMI albums are more responsive to the removal of DRM, then my estimates of the average effect of DRM removal would overstate the true effect. Given that my baseline estimates control for album and month fixed effects, a confounding omitted variable would need to be album-specific *and* time-varying.

Several parts of my empirical strategy help address this problem. First, I restrict my sample to albums released by record labels three years before 2007 to mitigate the concern that album release is influenced by DRM removal. For example, I assume the decision to release an album in the 1990s does not anticipate EMI's decision to drop DRM in 2007. In other words, the sample consists of albums where the decision to drop DRM is not correlated with ex ante album characteristics. Second, I apply a stringent matching procedure employing "coarsened exact matching" (CEM) (Iacus, King, and Porro, 2010), which includes pre-treatment online and offline sales,[20] the release year, and genre. Third, I include the monthly physical sales of the album as a control. This measure of offline album sales is album-specific and time-varying and allows the influence of offline sales and any correlates with offline sales (e.g., offline popularity, decade popularity) to vary over time. Fourth and

[20]Pre-treatment offline sales are from 1992-2006, while pre-treatment online sales are from 2003-2006.

importantly, the estimation on changes in EMI's sales distribution (i.e., top selling, lower-selling) after DRM removal only relies on variation within EMI's catalogue of albums. The core identifying assumption is that for each individual album released before 2007, EMI's decision to drop DRM is exogenous. In other words, the decision to drop DRM was not driven by any particular album but for the portfolio as a whole. Furthermore, recall from the institutional evidence above that EMI did not make the decision to drop DRM in anticipation of disproportionate changes in sales to their long tail catalogue. Thus, the exogeneity of the DRM "experiment" is more compelling for analyses that focus on the impact of DRM removal on changes in different parts of EMI's sales distribution compared to estimation on the overall effect of DRM removal on album sales.

5 Data

5.1 Data Construction and Sources

The primary data source for this study comes from Nielsen SoundScan, which is the principal source of sales data for the industry and the source of the well-known Billboard music charts. SoundScan tracks music sales at the point of sale, essentially monitoring cash registers at more than 14,000 retail, mass-merchant, and online stores in the United States. I also consult various other websites for auxiliary information (e.g., about genres and record labels) and to verify album release dates.[21]

My data covers music sold between January 1992 and June 2012. This dataset contains monthly data of the number of albums sold through traditional outlets like retail chains and also the sale of digital albums. Additionally, I have the monthly sales of digital tracks, which are songs purchased individually through online platforms like iTunes. SoundScan started tracking online sales in July 2003. Thus, for each album, I can calculate the number of units sold through traditional "brick and mortar" channels as well as through online channels. The sample covers the main Billboard genres: Pop/Rock, Country, Christian, Hip Hop and

[21]The primary source for genre information is http://www.allmusic.com/.

R&B, and Jazz & Classical. The unit of analysis is album-month.

In order to examine the effect of DRM removal on the entire distribution of music, it is important to collect a sample that is representative of the entire universe of available music. Given that music sales are highly skewed with a small number of artists responsible for the majority of music sales, only focusing on albums listed on the Top Billboard charts will not fully capture the effect of DRM removal on the entire body of the music sales distribution. Thus, I need to collect a sample of data that represents the hits (right tail), the middle, and the lower-selling albums (i.e., the long tail) of the sales distribution, which is not a trivial task.[22]

I begin my data collection process by collecting a list of all record labels currently operated by the four major record companies. While this seems relatively straightforward, there are numerous labels operating under each major record company, and many labels have become defunct or absorbed into other labels over the years. For example, Sony owns over thirty labels.[23] Further complicating matters, SoundScan does not report the major record company but rather the label under which the album is released. Thus, I manually matched each label to the major record company by consulting auxiliary sources. I identify a total of 145 labels operating under the four majors. Next, I try to collect a comprehensive list of artists that are under each label. First, I identify 4,063 unique artists signed across the four major record labels. I randomly select 900 artists from the list for my data collection. Then, I collect the complete discography for these artists. The advantage of collecting complete discographies for each artist (depth) rather than collecting more artists (width) is that I can track the entire evolution of their careers and any changes in label affiliations. This is particularly important for identifying the product discovery mechanism because it allows me to examine the impact of DRM removal on the sales of non-EMI albums of EMI artists, which

[22]It is important to point out that while my sample consists of a representative distribution of major-label music, it may understate the long tail if we also consider music released by indie labels. I do not include music released by indie labels in the sample because there is no experiment that allows me to clearly identify fluctuations in sales due to changes in IP strategy. Furthermore, indie labels have been typically the first to abandon DRM. Nevertheless, given that major labels control more than 80% of the distribution of the industry, my sample captures a distribution of music that is commercially relevant.

[23]Typically, majors have different labels for different genres. For example, Sony has several labels under its Sony Music Nashville branch to oversee their country artists.

holds changes to relative price and piracy constant. After eliminating albums released after 2007 (i.e., the treatment date), the final sample consists of 5,864 albums from 634 artists. To my knowledge, this is the largest random sample of albums collected for an empirical study.[24]

5.2 Descriptive Statistics

Table 1 reports variable definitions and summary statistics for the sample. Sony, Warner, and Universal account for more than 26% of the sample, while EMI accounts for 17% of the sample. The sample covers a wide range of album vintages and includes albums released between 1975 to 2006.

Like the sale of most creative products, such as books and movies, music sales are extremely skewed, particularly for *Physical* album (i.e., WalMart) sales. An average album sells around 3,440 copies per month through brick and mortar channels, although there are albums that sell zero copies a month to more than 3 million copies a month. *Online* albums sell on average 330 copies per month. The variance around online sales is noticeably smaller, suggesting that online sales are perhaps less dominated by hit albums.

Under 10% of the sample sells more than 1 million copies in the first three years of release, whereas over 48% of the sample sell less than 50,000 copies during the first three years of release. That is, the number of copies sold by the top 10% is at least 20 times more than what is sold by almost half of the albums in the sample. Looking at the fraction of total sales, albums that sell more than 1 million copies make up over 69% of total sales, while albums that sell fewer than 50,000 copies make up 2% of total sales.

The key explanatory variables are EMI and $PostDRMRemoval$. I use the indicator EMI to identify albums released by EMI. In other words, I use EMI to distinguish between my "treated" versus "control" albums. $PostDRMRemoval$ is equal to 1 after EMI drops DRM in April 2007 for all albums to capture counterfactual changes in album sales had other record companies dropped DRM at that time.

[24]Oberholzer-Gee and Strumpf (2007) sample 680 albums from Billboard charts.

6 Results

In the sections below, I start by estimating the impact of DRM removal on EMI's album sales using the original sample and a matched sample, where each treated album is paired with "similar" control albums. I examine whether the main results are driven by endogenous timing by investigating the presence of pre-trends. Then I investigate whether dropping DRM facilitates product discovery in the long tail by examining whether DRM removal disproportionately impacts different parts of EMI's sales distribution, followed by albums of different vintages and genres.

6.1 Main Results

Table 2 estimates the effect of removing DRM on EMI's online album sales. All baseline specifications include album fixed effects, which accounts for heterogeneity in the underlying quality of individual albums, such as total album sales, vintage, and genre, and month-year fixed effects, which control for album-invariant changes over time.

Column (1) implements my main specification in Equation (1).[25] To disentangle the marginal impact of DRM removal from the selection effect, I develop a difference-in-differences estimator that identifies the average differences in digital album sales between the treated (i.e., EMI albums) and control albums (i.e., non-EMI albums) and the change in sales that results from DRM removal. Since all albums eventually drop DRM by April 2009, I trim the sample at April 2009. $EMI \times PostDRMRemoval$ is a dummy variable equal to 1 only in those years after DRM is removed from EMI's albums. I find the marginal impact of DRM removal on album sales is around 13%.

Column (2) excludes holiday and compilation albums in order to focus on unique studio albums that are not affected by seasonality. I find that the marginal impact of DRM removal on EMI's album sales falls to around 8%. Column (3) includes a polynomial time trend of degree six to flexibly control for differences in sales due to album release date.

[25]Note that EMI_i is not separately included in the regression because I identify it through the album-fixed effects.

Even though I include album and month-year fixed effects, one may be concerned that there are confounding album-specific and time-varying characteristics that affect sales. For instance, DRM removal may have a larger impact on popular albums, even though popularity can hinder album sales because popular albums are actively pirated. To address this possibility, I include logged monthly physical sales (i.e., offline sales) in Column (4) as a control. This album-specific and time-varying measure of album popularity allows the influence of offline sales and any correlates with offline sales to vary fully flexibly over time. Including this measure actually increases the estimated effect of DRM removal on album sales to 10%.[26]

6.2 Robustness

So far my analysis has assumed that the timing of the removal of DRM is uncorrelated with factors that determine the outcomes of interest, conditional on the baseline controls. If EMI albums experience a significant increase in digital album sales prior to DRM removal, this would imply that the measured post-DRM effect is confounded with a pre-DRM removal trend, undermining the effect of β as a treatment effect. If there is no pre-trend, then one would be more convinced that the main results are not driven by endogenous timing. To investigate the presence of pre-trends, I estimate the following equation:

$$log(OnlineAlbumSales+1)_{it} = \alpha + \sum_{k=-7}^{7} \beta_k (EMI_i \times \mu_t) + \sum_{s=1}^{6} w_s t^s + X_{it} + \delta_i + \mu_t + \epsilon_{it}, \quad (2)$$

where I interact the treatment variable with a series of dummy variables for each quarter preceding and following DRM removal, along with the album-specific and time-varying measure of popularity, album, and quarter-year fixed effects and a polynomial time trend of

[26] This estimate is very close to the actual increase in sales experienced by EMI. Barney Wragg, EMI's Chief of Digital Operations stated: "At the time when we removed DRM from our downloads at EMI, we saw about a 10% increase in our digital sales".

degree six. For this estimate, I group all observations that are eight or more quarters prior together. This is the reference group that is omitted.

Figure 1 plots each of these estimates, where each point on the graph represents the estimated difference between EMI and non-EMI online album sales in that quarter. Two findings stand out. First, although the pre-DRM removal album sales pattern suggests that the average quarterly difference in sales between EMI and non-EMI albums is around 22%, the pre-DRM removal does not suggest a clear upward trend in the years prior to DRM removal. Second, the sizable increase in sales in the months following DRM removal is consistent with DRM removal having a significant impact on album sales. While EMI albums experience a 30% increase in the quarter immediately following DRM removal, this effect increases to roughly 43% one year later. Given that the average quarterly difference between EMI and non-EMI album sales before DRM removal is roughly 22% and the average difference between EMI and non-EMI albums in the period after DRM removal is 32%, DRM removal boost EMI's album sales by around 10%, which is roughly equal to the estimated coefficient in Table 2.

Another concern when estimating the effect of DRM removal on album sales is the emergence of the Amazon online store in October 2007. Recall that EMI drops DRM in April 2007 while the other major record companies do not drop DRM completely until April 2009. Given that Amazon carries a selection of the record company's DRM-free content, one might be concerned that the observed increase in EMI's album sales can be attributed to an additional outlet for selling its content rather than DRM removal. This concern is alleviated through the inclusion of month-year fixed effects, which capture any time-varying changes during this period. Furthermore, the fact that there is a large increase before October 2007 in Figure 1 shows that the boost in sales is not driven by the availability of DRM-free tracks on Amazon.

To alleviate concerns that treated and control albums are not similar, I construct a matched sample using a stringent matching procedure called Coarsened Exact Matching (CEM). I match albums on total online and physical sales of the album before DRM removal, release year, and genre. As the descriptive statistics reported in Table 3 demonstrate, the treated (EMI) and control (non-EMI) albums in this sample are indeed more similar than in the original sample. Table 4 replicates the same regression results from Table 2 using the matched CEM sample. The results are largely consistent across both samples. In fact, Column (4) suggests a slightly larger increase in online album sales after DRM removal (13%) compared to the results from the unmatched sample (10%). Given that the matching methodology relies on the same conditional distribution as the original sample (Angrist and Pischke, 2009) and that I lose observations through the matching procedure, I proceed with the original sample in the rest of my results below, although they are robust to the matched sample.

6.3 The Long Tail Effect and Product Discovery

DRM removal substantially lowers the cost of sharing legally purchased digital music, which has two countervailing effects on online sales. On the one hand, removing sharing restrictions may decrease the cost of piracy. To the extent that pirated copies are substitutes for the original copy and the pirated copies are songs with valuations above the price, then relaxing sharing restrictions may lead to lower online sales.[27] On the other hand, relaxing sharing restrictions may increase online sales by facilitating product discovery.

The argument for product discovery is as follows. The volume of music that is commercially available is vast and spans many genres and artists. Given the large volume of available music, consumers do not have all of the information on all music items. Further, music is an experience good whose value is revealed to the consumer after initial consumption. Sampling allows consumers to gain information about product fit before purchase (Chellappa and Shivendu, 2005; Peitz and Waelbroeck, 2006).

Consider a world where there are two types of artists: popular and unpopular. Popular artists are defined as those artists whose music is known by a greater fraction of the population. Conversely, unpopular artists are those artists whose music is known by a smaller

[27] If the pirated copies are songs with valuations below the price - and would otherwise not have been purchased - then file sharing raises consumer welfare without reducing industry revenue.

fraction of the population. Thus, information on unpopular artists is harder to find.

DRM removal relaxes sharing restrictions, which means it is easier to share DRM-free music with family and friends and upload them onto file sharing networks. Thus, DRM removal lowers consumers' sampling and search costs. DRM removal allows consumers to determine the true value of a music item from an unpopular artist in two ways. First, unpopular music is more likely to be found on file sharing networks after DRM removal because they are more easily shared. Second, sharing from family and friends is often an efficient form of word-of-mouth advertising because it allows consumers to prioritize information, especially in environments where consumer attention is scarce (Gans, 2012). Moreover, sharing information can also generate networks effects, where joint consumption is more valuable than individual consumption (Takeyama, 1994; Gayer and Shy, 2003). I am not able to distinguish between these two channels empirically but consider the overall impact of relaxing sharing restrictions on different parts of the sales distribution.

However, the extent to which DRM removal impacts music sales is likely to be different for popular and unpopular artists. The reason is that consumers are unlikely to sample and subsequently buy the popular artist's music as a result of DRM removal because they have already experienced the music beforehand (e.g., through radio). Thus, relaxing sharing restrictions is unlikely to increase the sales of music from popular artists. In contrast, DRM removal facilitates sampling of music from unpopular artists that otherwise would not have occurred, which can subsequently lead to the purchase of other music by the unpopular artist that otherwise would not have been purchased in the absence of DRM removal.

Indeed, Gopal, Bhattacharjee, and Sanders (2006) show that as the cost of sampling goes to zero, consumer surplus is maximized by the consumption of either popular or unpopular music (if the true values of popular and unpopular artists' music are equal) and hence the difference in sales between a popular artist's music and that of an unpopular artist becomes negligible. Note that relaxing sharing restrictions will not affect the set of consumers who would buy in the absence of sharing technologies nor the set of consumers that would always pirate regardless of search costs.

Next, I examine whether the removal of DRM disproportionately benefits albums at different parts of the sales distribution. In Table 5, I define the sales distribution based on the total number of albums sold in the first three years after album release. I also limit the sample to albums released before 2004 and run the estimation on the period after 2004. Columns (1) and (2) estimate the marginal impact of DRM removal on EMI albums that have sold more than 1 million and 500,000 copies, respectively. These are albums designated as "Platinum" and "Gold," respectively by the RIAA.[28] Columns (3) and (4) estimate the marginal impact of DRM removal on EMI's album sales for the middle part of the sales distribution. I present two definitions of the "middle": albums that have sold between 100,000 and 1 million copies (Column 3) and albums that have sold between 50,000 and 500,000 copies (Column 4). Columns 5-7 focus on the long tail - the lower-selling albums in the distribution. I define the tail in three ways: albums that have sold less than 100,000 copies (Column 5), albums that have sold less than 50,000 copies (Column 6), and albums that have sold less than 25,000 copies (Column 7).

I find that the marginal impact of DRM removal on the top-selling albums is negative and insignificant, which suggest that DRM removal does not appear to benefit top-selling albums. While the coefficients for the top-selling albums are insignificant, this does not mean that we can conclude DRM removal had no effect on top-selling albums. For instance, the 95% confidence interval for the coefficient in Column (1) is quite wide - between -0.39 and 0.21, which means that it is possible that the true impact of DRM removal on top-selling albums is large and negative. In other words, removing sharing restrictions for popular albums can lead to reductions in sales, perhaps by making it less costly for consumers to engage in piracy. On the other hand, the true impact of DRM removal on top-selling albums may be positive. Similarly, I find that the impact of DRM removal on the middle part of the sales distribution is negative but statistically insignificant.[29]

[28] http://www.riaa.com/goldandplatinum.php?content_selector=new-combined-GP [accessed June 2, 2013].

[29] Bar-Issac, Caruana, and Cunat (2010) show that lower search costs can simultaneously account for both superstar and long tail effects, with sales to both the head and tail of the sales distribution coming from middling firms whose designs change from broad to niche strategy.

In contrast, I find that the impact of DRM removal on EMI's least popular albums is positive and significant across all three definitions of the long tail. Interestingly, the magnitude increases for albums further down the tail - DRM removal increases EMI's tail albums that sell less than 25,000 copies by 30%, compared to the 24% increase for the tail albums that sell less than 100,000 copies. Figure 2 plots pre- and post-DRM removal effects on the long tail albums and shows that EMI's long tail albums experienced an almost 30% increase in the quarters after DRM removal relative to non-EMI albums.[30]

Note that my results do not imply that sales from discovery offsets piracy. If total consumption increased after DRM removal, it is possible that piracy may have increased more than sales. For instance, DRM removal may facilitate file sharing from consumers who otherwise would not have purchased the tail music. In this case, DRM removal raises consumer welfare without reducing firm revenue. My results indicate that EMI's net revenue increased irrespective of potential changes in piracy due to DRM removal.

These results are consistent with theory that suggests lowering search costs can increase the sale of niche products (Bar-Issac, Caruana, and Cunat, 2010; Yang, 2012) and sharing can facilitate product discovery (Peitz and Waelbroeck, 2006; Gans, 2012). Given that the average album in the tail (below 100,000 copies) sells around 28,000 copies in the first three years of release, DRM removal boosts album sales by over 7,500 copies on average. Taking the point estimates of my results, some simple back-of-the-envelope calculations reveal that EMI sells an additional 815,512 copies after DRM removal.[31] Given that price stays relatively constant during this period and assuming each album costs on average \$10, dropping DRM boosted EMI's revenues by over \$8 million.

[30]I explore alternative ways of defining different parts of the distribution in the Appendix. For example, in Table A1, I explore whether the main result holds using triple interactions. In Table A2, I repeat results in Table 5 using the matched CEM sample. The results are largely consistent. In Table A3, I define the sales distribution based on a rank ordering of albums. For the sake of comparison, albums below the top 2000 albums sell on average 15,712 copies in the first three years, compared to the 3,835,923 copies sold by the top 200 albums. My results are largely consistent.

[31]Taking the total online sales in each part of the distribution and the point estimates from Table 5: = 2978892*0.27+1245703*0.009 = 815,512.

6.4 The Long-Tail Effect and Increase in Value

An alternative mechanism for the observed increase in the sales of albums in the long tail is that dropping DRM increases the value of the album because consumers can now use the music in more ways. In other words, consumers may purchase more music at the tail not necessarily because they have discovered new music due to relaxed sharing restrictions but rather because they are more likely to consume music in the long tail since the relative cost of purchasing less popular music is lower. However, given that DRM has dropped the relative price for all music, it is unlikely that consumers are completely price inelastic for the top-selling and mediocre albums. In other words, it is unlikely that consumers will not purchase more top-selling hits given a drop in price if we assume constant price elasticity across the sales distribution. Since I do not find a statistically significant effect of DRM removal on albums that are in the top or middle part of the sales distribution, it is unlikely that the increase in sales of lower-selling albums is largely driven by a drop in relative price.

Nevertheless, I examine this possibility by considering the impact of DRM on the sales of non-EMI albums of EMI artists (Table 6). Consider an EMI artist, like Al Green, who has released albums with other major record companies (e.g., Universal) in the past. If the product discovery mechanism holds, then dropping DRM on the artist's EMI albums should also lead to an increase in the sale of its back catalogue of non-EMI albums, even though their relative price has not fallen. Indeed, I find that DRM removal increases the sale of EMI artists' non-EMI albums in the long tail but does not benefit their top-selling non-EMI albums. Not surprisingly, the magnitude of increase for non-EMI albums of EMI artists is smaller than the impact on EMI albums of EMI artists. This is a compelling piece of evidence for the product discovery mechanism; by looking at non-EMI albums that are not impacted by DRM removal for the same artist, I am holding changes to relative price and piracy constant.

6.5 The Long Tail Effect: Albums of Different Vintages & Genres

Chris Anderson (2004) famously said, *"You can find everything out there on the Long Tail. There's the back catalog, older albums still fondly remembered by longtime fans or rediscovered by new ones ... niches by the thousands."* In this section, I consider the impact of DRM removal on different types of albums in the long tail that are most likely to benefit from product discovery: older albums and albums of different genres.

Tables 7 and 8 examine the effect of DRM removal on old and new albums, respectively. I define old albums as albums released before 1992, although results are robust to other age cutoffs. If older albums are more likely to fall in the long tail and are difficult to discover, I should find a larger effect for albums of older vintages compared to newer albums. Another reason for separately looking at albums released before 1992 is because SoundScan started tracking album sales in 1992. Thus, for albums released before 1992, I can only capture total sales in the first three years that SoundScan started tracking its sales, which understates the three-year total sales of albums released before 1992.

Table 7 examines the impact of DRM removal for old albums (i.e., albums released before 1992). Column (1) shows that the effect of DRM removal on EMI's online sales for old albums is 23%. However, consistent with results in Table 5, only albums in the long tail benefit from DRM removal. Column (6) shows that albums that sell less than 100,000 copies experience an increase of 29%, while albums that sell less than 25,000 copies (Column 8) have a boost in sales of around 41%. There is not a statistically significant effect on albums that fall in the top or middle part of the sales distribution.

Table 8 examines the impact of DRM removal for new albums (i.e., albums released after 1992). Several interesting results emerge. First, I find that the overall increase in online sales is small and statistically insignificant, in contrast to old albums (23%). Second, consistent with results on old albums in Table 7, DRM removal increases the online sale of albums at the tail of the distribution by around 21% - 25%. While the size of these coefficients are smaller than the coefficients for the tail of older albums, the difference in size is statistically insignificant. While the overall impact on sales is larger for older albums compared to newer albums, these results suggest that product discovery is more likely to disproportionately benefit less popular albums at the tail of the distribution regardless of vintage. In other words, tail albums of older vintages do not appear to significantly benefit more from relaxed sharing restrictions compared to tail albums of newer vintages.

Next, I consider the impact of DRM removal on music of different genres. Specifically, I compare hip hop and R&B to jazz and classical albums (Table 9). Anecdotal evidence suggests that hip hop and R&B are the most pirated genres, while jazz and classical are the least pirated genres.[32] Two interesting results stand out. First, I find the overall change in sales for hip hop and R&B is small and statistically insignificant whereas DRM removal increases jazz & classical music sales by 28% overall. Second, I do not find evidence that DRM removal significantly impacts any part of the sales distribution of hip hop albums whereas DRM removal increases the sales of less popular jazz and classical music by at least 31%.

One alternative explanation may be that there is heterogeneity in unobserved user distaste for the inconvenience of DRM which are correlated with unobserved heterogeneity for demand for music in a particular segment. For instance, if the average jazz and classical buyer owns a large number of devices due to unobserved higher income than the average hip hop buyer, then removing DRM would remove a distasteful attribute for which the first group would be willing to pay more than the second group, and consequently would generate more sales from jazz and classical buyers. While this may explain the overall difference between the genre cohorts, it is unlikely that the increase in sales of jazz and classical albums is concentrated only in lower-selling albums. If removing DRM is only removing distaste, we should also expect to see an increase in top-selling jazz and classical albums as well.[33] Similarly, while

[32]http://www.makeuseof.com/tag/top-10-pirated-music-bittorrent-today/ [accessed May 5, 2013]. This is also confirmed from descriptive evidence in Oberholzer-Gee and Strumpf (2007) and from EMI's surveys (Danaher, Smith, Telang, and Chen, 2013).

[33]It is also unclear whether hip hop and rap listeners own fewer devices compared to jazz and classical consumers. Survey evidence shows that the average age of rap and hip hop listeners is younger than the average jazz and classical listener (SLMG, 2004; NEA, 2008), and younger Americans are more likely to own more devices (see Table A4). Thus, while total income may be higher for jazz and classical consumers, it is not obvious that they would own more devices (and consequently have greater distaste for DRM) compared to hip hop and R&B consumers.

it is plausible that listeners of jazz and classical music are less likely to engage in piracy, it is unlikely that the increase in sales after DRM removal is concentrated only in lower-selling albums while the impact on the top and middle part of the distribution is insignificant. Taken together, this may suggest that certain demographics benefit more from discovery as a result of relaxing sharing restrictions. For instance, the hip hop demographic may not be as inhibited by file sharing and thus are not strongly impacted by DRM removal.

Overall, my results are consistent with the long tail hypothesis, which predicts a shift in consumption away from hits to a much larger number of lower-selling niche products provided through online channels. DRM removal disproportionately benefits poorer-selling albums, such as albums in the back catalogue and niche genres. While the long-tail literature argues that these changes are largely due to supply-side changes, such as lower distribution costs, the DRM shock adds nuance to this story by providing a consumption-based argument.

DRM removal substantially lowers the cost to sharing legally purchased digital music, which may decrease sales by facilitating piracy and increase sales by facilitating product discovery, in particular for less popular music. My results suggest that for popular music, the net change in sales from relaxed sharing restrictions is small, likely because it is already discovered and pirated before DRM removal. In contrast, DRM removal facilitates sharing of music from unpopular artists that otherwise would not have occured, which can subsequently lead to the purchase of other music by the unpopular artist. After all, sharing from the right people (i.e., friends and family) prioritizes information and facilitates better matches (Gans, 2012).[34] This result is also consistent with the word-of-mouth literature (Dellarocas, 2003; Godes and Mayzlin, 2004) that finds consumers rely on word-of-mouth for riskier transactions, such as niche products.

[34]Tucker and Zhang (2011) show that popularity information can benefit niche products disproportionately in online markets.

7 Conclusion and Implications

Digitization has materially lowered the costs of production and distribution and increased the variety of products available for consumption in many industries. The economic consequences go far beyond a decline in costs. Digitization has initiated significant shifts in market structure and changes in competitive behavior in many media markets and has been closely associated with ushering in Schumpeterian creative destruction in many knowledge-based industries (Greenstein, 2010). In these settings, how should firms design their IP strategy to balance the incentives for product discovery with the incentives for legal purchases in the digital economy?

DRM is a prevalent strategy implemented by firms in media industries (e.g., books, movies, and video games) that highlights this tension. DRM is a unique counter-piracy measure because it is a strategy that is implemented by firms rather than by IP policy and law enforcement.[35] Specifically, DRM's sharing restrictions have countervailing effects on sales. While it has the potential to combat piracy, it may also hinder product discovery, both of which are salient issues in many digital markets. Thus, the recorded music industry removing DRM on music at different times provides the first empirical evidence of a more "relaxed" digital copyright strategy on digital sales and its heterogeneous effects on different parts of the sales distribution.

My analysis in this paper, based on a large representative sample of albums from all four major record companies, sheds light on this question. I find that the removal of DRM increases digital sales by 10%. More importantly, the effect is most pronounced for albums at the long tail of the music sales distribution, providing support for the long tail hypothesis that lowering search costs can facilitate product discovery of non-mainstream fare.

My results indicate that some firms in creative industries may optimally choose a relaxed IP strategy given the composition of their sales distribution and an enforceable legal frame-

[35] An analogous response by the book publishing industry is to hire private companies to protect book titles from piracy. Reimers (2014) finds that this form of piracy protection increases the sale of e-books, but the protection is most effective for popular titles.

work.[36] My results also suggest that firms in these settings need to consider IP strategy as part of their broader product market strategy, since it is tied to sales in many contexts. Given the policy debates surrounding fair use (Lichtman, 2009), my results also suggest that expanding fair use in a similar way would arguably benefit consumers and those who want to remix, etc.,[37] and have small harm on certain types of copyright holders. Note that this is not a welfare analysis of the consequences of DRM removal as piracy data during this period is unavailable. It is possible that piracy may have increased during this period. For instance, if DRM removal facilitated file sharing from consumers who otherwise would not have purchased long tail music, DRM removal raises consumer welfare without reducing firm revenue. My results indicate that relative to a regime of DRM, EMI's music sales increased on net and disproportionately for long tail content.

My analysis is of course subject to limitations such that generalizing to other contexts should be done with caution. Other settings, such as books, movies, and video games, are different from the recorded music industry in many respects. Notably, products in these other industries take a longer time to consume compared to listening to a song. Arguably, consumers also place different values on repeat consumption of books and movies. Furthermore, the discovery process for other creative goods is likely quite different from music, and there are likely fewer complementary ways to substitute for product consumption. For example, research shows that while file sharing has reduced physical sales, demand for live concerts has increased (Mortimer, Nosko, and Sorensen, 2012) and concert prices are sensitive to search cost reductions in secondary markets (Bennett, Seamans, and Zhu, 2013). It is difficult to identify whether there are similar complementary activities in other creative industries. My results are based on U.S. data and thus restricted to a setting where the legal framework for IP is enforceable relative to other settings where the appropriability regime is weaker. In settings where the legal framework is weakly enforced, it may not be optimal for

[36] Interestingly, DiCola (2013) shows through survey evidence that copyright only benefits the revenue of top musicians in the top income bracket but the vast majority of musicians do not depend on copyright for music revenue.

[37] See McLeod and DiCola (2011) for a discussion on how copyright may constrain musicians building on another's prior works.

firms to relax sharing restrictions, and they instead should consider alternative mechanisms to appropriate returns to innovation. Exploring the margins most influenced by digitization and the effect of IP strategies on the distribution of consumption and production patterns will continue to be a prominent line of inquiry for scholars of innovation and competition in the years ahead.

References

Agrawal, A., and R. Henderson (2002): "Putting Patents in Context: Exploring Knowledge Transfer from MIT," *Management Science*, 48, 44–60.

Aguiar, L., and J. Waldfogel (2014): "Quality Predictability and the Welfare Benefits from New Products: Evidence from the Digitization of Recorded Music," *Working Paper*.

Anderson, C. (2004): "The Long Tail," `http://www.wired.com/wired/archive/12.10/tail.html`, [Online accessed September 5, 2012].

——— (2006): *The Long Tail: Why the future of business is selling less of more*. Hyperion Books, New York.

Angrist, J., and J. Pischke (2009): *Mostly Harmless Econometrics: An Empiricists Companion*. Princeton University Press, Princeton, NJ.

Arrow, K. (1962): "Economic Welfare and the Allocation of Resources for Invention," in *The Rate and Direction of Inventive Activity: Economic and Social Factors*, ed. by R. Nelson. Princeton University Press, Princeton, NJ.

Bakos, Y., and E. Brynjolfsson (1999): "Bundling information goods: Pricing, profits and efficiency," *Management Science*, 45, 1613–1630.

Bar-Issac, H., G. Caruana, and V. Cunat (2010): "Search, Design, and Market Structure," *American Economic Review*, 102, 11401160.

Belleflamme, P., and M. Peitz (2011): "Digital Piracy: Theory," in *The Oxford Handbook of the Digital Economy*, ed. by M. Peitz, and J. Waldfogel. Oxford University Press.

Bennett, V., R. Seamans, and F. Zhu (2013): "Value Appropriation, Search Frictions, and Secondary Markets," Working Paper.

Bertrand, M., E. Duflo, and S. Mullainathan (2004): "How much should we trust differences-in-differences estimates?," *The Quarterly Journal of Economics*, 119, 249–275.

Bhattacharjee, S., R. Gopal, K. Lertwachara, J. Marsden, and R. Telang (2007): "The effect of digital sharing technologies on music markets: A survival analysis of albums on ranking charts," *Management Science*, 53, 1359–1374.

Blackburn, D. (2004): "Online Piracy and Recorded Music Sales," Mimeo.

Brynjolfsson, E., Y. Hu, and D. Simester (2011): "Goodbye Pareto Principle, Hello Long Tail," *Management Science*, 57, 1373–1386.

Brynjolfsson, E., Y. Hu, and M. Smith (2006): "From Niches to Riches: The anatomy of the long tail," *Sloan Management Review*, 47, 67–71.

Chellappa, R., and S. Shivendu (2005): “Managing Piracy: Pricing and Sampling Strategies for Digital Experience Goods in Vertically Segmented Markets,” *Information Systems Research*, 16, 400–417.

Danaher, B., M. Smith, T. Telang, and S. Chen (2013): “The Effect of Graduated Response Anti-Piracy Laws on Music Sales: Evidence from an Event Study in France,” *Journal of Industrial Economics*, Forthcoming.

Dellarocas, C. (2003): “The digitization of word of mouth: Promises and challenges of online feedback mechanisms,” *Management Science*, 49, 1407–1424.

Dewan, S., and J. Ramaprasad (2012): “Music blogging, online sampling, and the long tail,” *Information Systems Research*, 23, 1056–1067.

DiCola, P. (2013): “Money from Music: Survey Evidence on Musicians’ Revenue and Lessons About Copyright Incentives,” *Arizona Law Review*, pp. 1–28, Forthcoming.

Elberse, A., and F. Oberholzer-Gee (2007): “Superstars and underdogs: An examination of the long tail phenomenon in video sales,” Harvard Business School Working Paper Series, No. 07-015.

Ellison, G., and S. Ellison (2006): “Internet retail demand: Taxes, geography, and online-offline competition,” NBER Working Paper 12242, National Bureau of Economic Research, Cambridge, MA.

Forman, C., A. Ghose, and A. Goldfarb (2009): “Competition Between Local and Electronic Markets: How the Benefit of Buying Online Depends on Where You Live,” *Management Science*, 55(1), 47–57.

Galasso, A., and M. Schankerman (2013): “Patents and Cumulative Innovation: Causal Evidence from the Courts,” Working Paper.

GALLUP (2013): “Americans’ Tech Tastes Change With Times,” `http://www.gallup.com/poll/166745/americans-tech-tastes-change-times.aspx`, [Online accessed June 5, 2014].

Gans, J. (2012): *Information Wants to be Shared*. Harvard Business Review Press, Boston, MA.

Gayer, A., and O. Shy (2003): “Internet and Peer-to-Peer Distribution in Markets for Digital Products,” *Economic Letters*, 81, 51–57.

Godes, D., and D. Mayzlin (2004): “Using online conversation to study word of mouth communication,” *Marketing Science*, 23, 545–560.

Goolsbee, A. (2001): “Competition in the computer industry: Online versus retail,” *Journal of Industrial Economics*, 49, 487499.

Gopal, R., S. Bhattacharjee, and G. Sanders (2006): "Do Artists Benefit from Online Music Sharing?," *The Journal of Business*, 79, 1503–1533.

Greenstein, S. (2010): "The economics of digitzation: An agenda," Mimeo.

Hendricks, K., and A. Sorensen (2009): "Information and the skewness of music sales," *Journal of Political Economy*, 117, 324–369.

Iacus, S., G. King, and G. Porro (2010): "Causal inference without balance checking: Coarsened exact matching," Working paper, Harvard University.

International Federation of Phonographic Industry (2005): "IFPI releases denitive statistics on global market for recorded music," `http://www.ifpi.org/sitecontent/publications/rin_order.html`, [(August 2) London, UK].

Jaffe, A., M. Trajtenberg, and R. Henderson (1993): "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations," *The Quarterly Journal of Economics*, 108(3), 577–598.

Li, X., M. Macgarvie, and P. Moser (2013): "Dead Poets' Property - Do Stronger Copyrights Increase the Price of Books?," Working Paper.

Lichtman, D. (2006): "Defusing DRM," *Chicago Working Papers in Law Economics No.282*, pp. 1–10.

——— (2009): "Copyright as Innovation Policy: Google Book Search from a Law and Economics Perspective," in *Innovation Policy and the Economy*, ed. by J. Lerner, and S. Stern, vol. 9, pp. 57–80. University of Chicago Press.

Liebowitz, S. J. (2006): "File Sharing: Creative Destruction or Just Plain Destruction?," *Journal of Law and Economics*, 49(1), 1–28.

McLeod, K., and P. DiCola (2011): *Creative License: The Law and Culture of Digital Sampling*. Duke University Press.

Mortimer, J., C. Nosko, and A. Sorensen (2012): "Supply responses to digital distribution: Recorded music and live performances," *Information Economics and Policy*, 24, 3–14.

Moser, P. (2013): "Patents and Innovation: Evidence from Economic History," *Journal of Economic Perspectives*, 27, 23–44.

Murray, F., and S. Stern (2007): "Do formal intellectual property rights hinder the free flow of scientific knowledge? An empirical test of the anti-commons hypothesis," *Journal of Economic Behavior and Organizations*, 63, 648–687.

NEA (2008): "Age and arts participation: A case against demographic destiny," `http://arts.gov/sites/default/files/2008-SPPA-Age.pdf`, [Online accessed June 5, 2014].

Nelson, R. (1959): "The simple economics of basic scientific research," *Journal of Political Economy*, 67, 297–306.

Oberholzer-Gee, F., and K. Strumpf (2007): "The effect of file sharing on record sales: An empirical analysis," *Journal of Political Economy*, 115, 1–42.

Peitz, M., and P. Waelbroeck (2006): "Why the Music Industry May Gain from Free Downloading - the Role of Sampling," *International Journal of Industrial Organization*, 24, 907–913.

Petrick, P. (2004): "Why DRM should be cause for concern: An economic and legal analysis of the effect of digital technology on the music industry," Working Paper No. 2004-09. Harvard Law School.

Png, I., and Q. Wang (2006): "Copyright Duration and the Supply of Creative Work: Evidence from the Movies," Working Paper, National University of Singapore.

Prasad, A., and V. Mahajan (2003): "How many pirates should a software firm tolerate? An analysis of piracy protection on the diffusion of software," *International Journal of Research in Marketing*, 20, 337–353.

Prince, J. (2007): "The beginning of online/retail competition and its origins: An application to personal computers," *International Journal of Industrial Organization*, 25, 139156.

Reimers, I. (2014): "Can Private Piracy Protection be Effective? Evidence from Book Publishing," Working Paper.

RIAA (2010): "Year End Shipment Statistics," `www.riaa.com/keystatistics.php`, [Online accessed December, 2012].

Rob, R., and J. Waldfogel (2006): "Piracy on the high C's: Music downloading, sales displacement, and social welfare in a sample of college students," *Journal of Law and Economics*, 49, 29–62.

Scherer, F. (2003): *Quarter notes and bank notes: The economics of music composition in the eighteenth and nineteenth centuries*. Princeton University Press, Princeton, NJ.

Schumpeter, J. (1934): *The Theory of Economic Development*. Harvard University Press, Cambridge, MA.

Sinha, R., F. Machado, and C. Sellman (2010): "Dont think twice, it's all right: Music piracy and pricing in a DRM-free environment," *Journal of Marketing*, 74, 40–54.

SLMG (2004): "Rapper's Delight: A Billion Dollar Industry," `http://www.nbcnews.com/id/4304261/#.U5dnkfmwJNN`, [Online accessed June 5, 2014].

Smith, M., and R. Telang (2009): "Competing with free: The impact of movie broadcasts on DVD sales and Internet piracy," *MIS Quarterly*, 33, 321–338.

Takeyama, L. (1994): "The Welfare Implications of Unauthorized Reproduction of Intellectual Property in the Presence of Network Externalities," *Journal of Industrial Economics*, 42, 155–166.

Tucker, E., and J. Zhang (2011): "How Does Popularity Information Affect Choices? A Field Experiment," *Management Science*, 57(5), 828–842.

Vernik, D., D. Purohit, and P. S. Desai (2011): "Music Downloads and the Flip Side of Digital Rights Management," *Marketing Science*, 42, 1–17.

Waldfogel, J. (2011): "Bye, Bye, American Pie: The Supply of New Recorded Music since Napster," NBER Working Paper No. 16882.

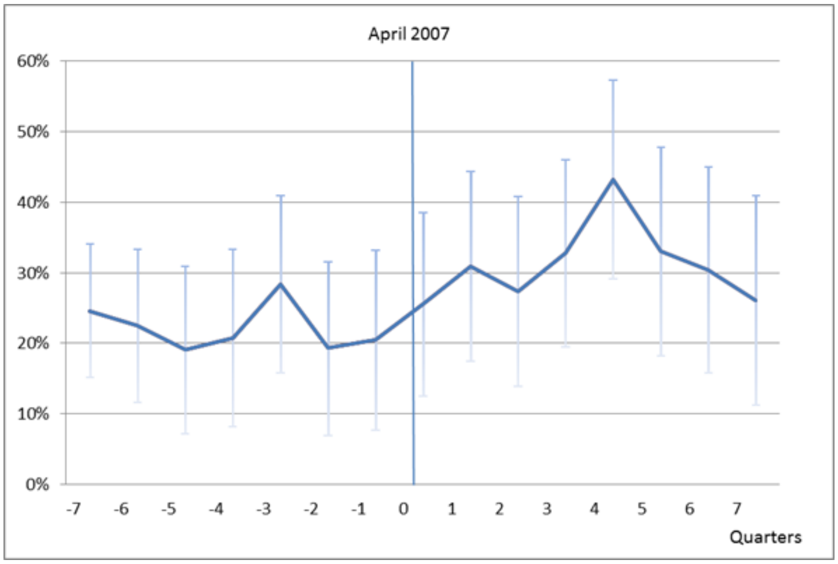
——— (2012): "Copyright Protection, Technological Change, and the Quality of New Products: Evidence from Recorded Music since Napster," *Journal of Law and Economics*.

Williams, H. (2013): "Intellectual Property Rights and Innovation: Evidence from the Human Genome," *Journal of Political Economy*, 121, 1–27.

Yang, H. (2012): "Targeted search and the long tail effect," Working Paper, Ohio State University.

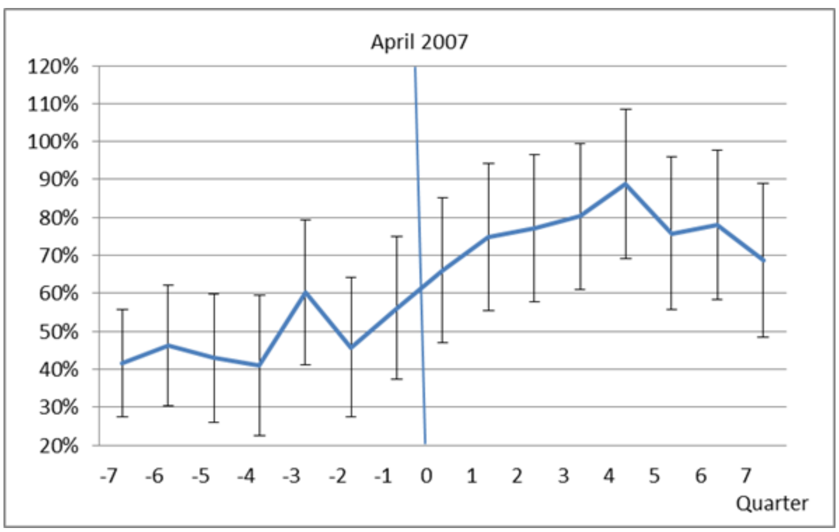
Zentner, A. (2006): "Measuring the effect of file sharing on music purchases," *Journal of Law and Economics*, 49, 63–90.

Figure 1: Pre- and Post DRM Removal Effects on Online Album Sales



Notes: Figure plots quarter by quarter pre-DRM removal and post-DRM removal changes to EMI's album sales from OLS regressions with dummy variables for each quarter preceding and following DRM-removal with album and quarter-year fixed effects, and a polynomial trend of degree six. Each point represents the estimated percentage difference between the treated (EMI) and control (non-EMI) albums sales in each quarter, along with upper and lower bounds for 95% confidence intervals.

Figure 2: Pre- and Post DRM Removal Effects on the Long Tail



Notes: Figure plots quarter by quarter pre-DRM removal and post-DRM removal changes to EMI's tail album sales (i.e. albums that have sold below 100,000 copies) from OLS regressions with dummy variables for each quarter preceding and following DRM-removal with album and quarter-year fixed effects, and a polynomial trend of degree six. Each point represents the estimated percentage difference between the treated (EMI) and control (non-EMI) albums sales in each quarter, along with upper and lower bounds for 95% confidence intervals.

Table 1: Summary Statistics

Variable Name	Description	Obs.	Mean	Std. Dev.	Min	Max
Sales Characteristic:						
Total Sales	The number of albums sold per month, including physical and online albums	709,594	3,708.176	28,355.69	0	3,348,623
Physical Sales	The number of offline albums sold per month (i.e., retail channels)	691,374	3,440.431	28,197.43	0	3,348,623
Online Sales	The number of track album equivalents sold per month (calculated by dividing total tracks sold per month by 10)	357,320	330.041	1,239.911	0	87,452.7
Total Sales - First three years after release	The number of total albums sold in the first three years of release (albums released before 2004)	136,861	351,118.9	933,383.3	12	12,758,683
Above 1 million copies	Equals 1 if the total albums sold in the first three years of release are under 1 million copies	12,008	2,624,900	1,962,438	1,001,699	12,758,683
Above 500k copies	Equals 1 if the total albums sold in the first three years of release are above 500,000 copies	21,251	1,789,186	1,758,543	501,956.4	12,758,683
Between 100k to 1 million copies	Equals 1 if the total albums sold in the first three years of release are between 100,000 and 1 million copies	41,487	342,618	226,924.3	100,003.5	995,057.9
Between 50k to 500k copies	Equals 1 if the total albums sold in the first three years of release are between 50,000 and 500,000 copies	49,531	180,778.4	120,088.3	50,062	499,823
Below 100k copies	Equals 1 if the total albums sold in the first three years of release are under 100,000 copies	83,366	27,834.95	26,304.45	12	99,995
Below 50k copies	Equals 1 if the total albums sold in the first three years of release are under 50,000 copies	66,079	16,319.07	13,327.47	12	49,976
Below 25k copies	Equals 1 if the total albums sold in the first three years of release are under 25,000 copies	49,395	9,797.895	7,212.879	12	24,994.4
Album Characteristic:						
Release Date	The release date of the album	709,594	1995.129	6.522682	1975	2006
Post-DRM-Removal	Equals 1 after April 2007 for all albums	709,594	0.392	0.488	0	1
EMI	The album's label is EMI	709,594	0.172	0.438	0	1
SONY	The album's label is Sony	709,594	0.262	0.44	0	1
WARNER	The album's label is Warner	709,594	0.266	0.442	0	1
UNIV	The album's label is Universal	709,594	0.299	0.458	0	1

Notes: The sample covers 5,864 albums from 634 artists. These albums cover all four major record labels (EMI, Sony, Warner, and Universal).

Table 2: The Impact of DRM Removal on Online Sales

Regression model: OLS	(1)	(2)	(3)	(4)
EMI x Post DRM Removal	0.126***	0.0882**	0.0911**	0.0990***
	(0.0338)	(0.0371)	(0.0365)	(0.0360)
log(Physical Sales)				0.108***
				(0.00928)
Exclude holiday & compilations		Yes	Yes	Yes
Polynomial time trend		Yes	Yes	Yes
Observations	357,320	254,749	254,749	254,749
R-squared	0.349	0.352	0.373	0.383
Number of albums	5,864	4,089	4,089	4,089

Notes: The dependent variable is logged online album sales. All specifications include album and month-year fixed effects.
Robust standard errors are clustered by album.
** p<0.01, ** p<0.05, * p<0.1

Table 3: Treated vs. Control Albums

Treated vs. Control albums in the ”Original Sample”

	Treated		Control	
	Mean	Std. dev.	Mean	Std. dev.
Pre-treatment online sales	3708	10607.44	7349	19574.02
Pre-treatment physical sales	337513	743631.3	455987	1117706
Release year	1995	6.144	1996	6.585

Treated vs. Control albums in the ”Matched Sample”

	Treated		Control	
	Mean	Std. dev.	Mean	Std. dev.
Pre-treatment online sales	4027	11022.5	4017	8071.983
Pre-treatment physical sales	308500	603877.2	249803	508900
Release year	1996	6.101	1995	6.517

Notes: This table compares the treated (i.e., EMI) and control (i.e., non-EMI) albums in the matched and un-matched sample. The CEM sample has fewer albums than the original sample because I drop the albums for which I find no match. The CEM procedure matches on pre-treatment online and offline sales, the release year, and genre.

Table 4: The Impact of DRM Removal on Online Sales (matched sample)

Regression model: OLS	(1)	(2)	(3)	(4)
EMI x Post DRM Removal	0.141***	0.101***	0.106***	0.115***
	(0.0342)	(0.0379)	(0.0373)	(0.0369)
log(Physical Sales)				0.101***
				(0.00980)
Exclude holiday & compilations		Yes	Yes	Yes
Polynomial time trend		Yes	Yes	Yes
Observations	316,408	225,392	225,392	225,392
R-squared	0.330	0.334	0.357	0.365
Number of albums	5,226	3,639	3,639	3,639

Notes: This table presents the same results as Table 2 using a matched sample based on Coarsened Exact Matching (CEM). All specifications include album and month-year fixed effects.
Robust standard errors are clustered by album.
** p<0.01, ** p<0.05, * p<0.1

Table 5: The Impact of DRM Removal on Different Parts of the Sales Distribution

Regression model: OLS	(1) Above 1M	(2) Above 500k	(3) 100k-1M	(4) 50k-500k	(5) Below 100k	(6) Below 50k	(7) Below 25k
EMI x Post DRM Removal	-0.0865	-0.00899	-0.0113	0.00870	0.214***	0.238***	0.264***
	(0.152)	(0.122)	(0.0711)	(0.0596)	(0.0484)	(0.0569)	(0.0574)
Observations	21,618	36,380	63,629	73,447	116,600	92,020	70,836
R-squared	0.359	0.316	0.272	0.253	0.154	0.145	0.145
Number of albums	346	581	1,027	1,194	1,932	1,530	1,177

Notes: I define different parts of the sales distribution based on the number of albums sold in the first three years of release for albums released before 2004. For example, Column (1) restricts the sample to albums that have sold over 1 million copies; Column (3) restricts the sample to albums that have sold between 100,000 and 1 million copies, and Column (6) restricts the sample to albums that have sold fewer than 100,000 copies. The regressions are run on the period after 2004. All specifications include album and month-year fixed effects. Polynomial time trend of degree six is included. Compilations and holiday albums are excluded.
Robust standard errors are clustered by album.
** p<0.01, ** p<0.05, * p<0.1

Table 6: The Impact of DRM Removal on non-EMI albums of EMI Artists Released Before DRM Removal

Regression model: OLS	(1) Overall	(2) Above 1M	(3) Above 500k	(4) 100k-1M	(5) 50k-500k	(6) Below 100k	(7) Below 50k	(8) Below 25k
$EMI_{EMIartist}$ x Post DRM removal	0.125***	-0.106	-0.041	-0.016	0.021	0.282***	0.308***	0.336***
	(0.044)	(0.156)	(0.127)	(0.076)	(0.064)	(0.055)	(0.065)	(0.065)
non-$EMI_{EMIartist}$ x Post DRM removal	0.004	0.083	-0.104	-0.210**	-0.133*	0.178***	0.239***	0.218***
	(0.042)	(0.246)	(0.171)	(0.084)	(0.068)	(0.047)	(0.052)	(0.058)
Observations	172,696	19,010	31,668	54,416	62,568	99,270	78,460	61,078
R-squared	0.188	0.352	0.301	0.259	0.244	0.144	0.137	0.136
Number of albums	2,826	304	506	880	1,018	1,642	1,302	1,011

Notes: $nonEMI_{EMIartist}$ is a new indicator variable that equals 1 for non-EMI albums of artists who have released an EMI album before DRM removal, and 0 otherwise. $EMI_{EMIartist}$ is an indicator variable that equals 1 for EMI albums of artists who have released an EMI album before DRM removal. I restrict the sample to albums released before 2004 and run the regression on the period 2004-2009. This table shows that while DRM removal does not significantly impact the overall sales of non-EMI albums released prior to DRM removal, it does disproportionately increase the sale of non-EMI albums in the lower end of the sales distribution. All specifications include a polynomial time trend of degree six and include album and month-year fixed effects.
Robust standard errors are clustered by album.
** p<0.01, ** p<0.05, * p<0.1

Table 7: The Impact of DRM Removal on Albums Released Before 1992

Regression model: OLS	(1) Overall	(2) Above 1M	(3) Above 500k	(4) 100k-1M	(5) 50k-500k	(6) Below 100k	(7) Below 50k	(8) Below 25k
EMI x Post DRM Removal	0.209***	0.134	0.379	0.037	0.097	0.255***	0.263***	0.345***
	(0.066)	(0.486)	(0.295)	(0.145)	(0.112)	(0.067)	(0.079)	(0.086)
Observations	73,529	2,931	5,465	6,354	26,073	52,677	41,991	32,219
R-squared	0.224	0.340	0.277	0.330	0.314	0.200	0.174	0.156
Number of albums	1,196	46	86	101	418	863	692	532

Notes: I define different parts of the sales distribution by the number of albums sold in the first three years after release. Note that because SoundScan starts recording sales in 1992, I calculate total sales based on the first three years in SoundScan's database. All specifications include album and month-year fixed effects and a polynomial time trend of degree six. Compilations and holiday albums are excluded. This table restricts the sample to albums released before 1992. Robust standard errors are clustered by album.
** p<0.01, ** p<0.05, * p<0.1

Table 8: The Impact of DRM Removal on Albums Released After 1992

Regression model: OLS	(1) Overall	(2) Above 1M	(3) Above 500k	(4) 100k-1M	(5) 50k-500k	(6) Below 100k	(7) Below 50k	(8) Below 25k
EMI x Post DRM Removal	0.047	-0.080	-0.044	-0.046	-0.028	0.189***	0.221***	0.214***
	(0.049)	(0.155)	(0.129)	(0.078)	(0.071)	(0.067)	(0.079)	(0.076)
Observations	129,561	18,934	31,199	46,149	47,875	64,478	50,487	38,935
R-squared	0.192	0.374	0.333	0.263	0.217	0.127	0.131	0.144
Number of albums	2,111	300	496	741	776	1,070	839	646

Notes: This table restricts the sample to albums released after 1992. Different parts of the sales distribution are defined by the number of albums sold in the first three years after release. All specifications include album and month-year fixed effects. Polynomial time trend of degree six is included. Compilations and holiday albums are excluded.
Robust standard errors are clustered by album.
** p<0.01, ** p<0.05, * p<0.1

Table 9: The Impact of DRM Removal on Albums of Different Genres

Hip Hop and R&B	(1) Overall	(2) Above 1M	(3) Above 500k	(4) 100k-1M	(5) 50k-500k	(6) Below 100k	(7) Below 50k	(8) Below 25k
EMI x Post DRM Removal	0.050	0.187	0.060	-0.169	-0.196	0.038	0.017	0.060
	(0.096)	(0.262)	(0.183)	(0.119)	(0.161)	(0.104)	(0.108)	(0.113)
Observations	33,179	6,409	10,906	12,246	9,990	14,524	12,283	10,194
R-squared	0.255	0.520	0.470	0.352	0.336	0.147	0.129	0.122
Number of albums	531	101	173	195	159	235	199	165

Jazz and Classical	(1) Overall	(2) Above 1M	(3) Above 500k	(4) 100k-1M	(5) 50k-500k	(6) Below 100k	(7) Below 50k	(8) Below 25k
EMI x Post DRM Removal	0.249***	-1.536**	-1.010	0.320**	0.125	0.327***	0.362***	0.273***
	(0.086)	(0.594)	(0.721)	(0.153)	(0.176)	(0.080)	(0.071)	(0.069)
Observations	35,855	384	1,073	5,925	10,161	29,546	24,621	19,744
R-squared	0.163	0.539	0.283	0.228	0.199	0.164	0.172	0.175
Number of albums	582	6	17	93	161	483	404	325

Notes: All specifications include album and month-year fixed effects. Polynomial time trend of degree six is included. Robust standard errors are clustered by album.
** p<0.01, ** p<0.05, * p<0.1

Appendix A Robustness

Table A1: The Impact of DRM Removal on Different Parts of the Sales Distribution (Triple Interactions)

Regression model: OLS	(1) Tail (Below 100k)	(2) Tail (Below 50k)	(3) Tail (Below 25k)
EMI x Post DRM Removal x Middle (100k-1M)	0.061		
	(0.178)		
EMI x Post DRM Removal x Tail (Below 100k)	0.314*		
	(0.170)		
EMI x Post DRM Removal x Middle (50k-500k)		0.055	
		(0.142)	
EMI x Post DRM Removal x Tail (Below 50k)		0.290**	
		(0.138)	
EMI x Post DRM Removal x Middle (25k-500k)			0.105
			(0.141)
EMI x Post DRM Removal x Tail (Below 25k)			0.289**
			(0.138)
EMI x Post DRM Removal	-0.092	-0.048	-0.048
	(0.162)	(0.126)	(0.126)
Post DRM Removal x MIddle (100k-1M)	-0.336***		
	(0.064)		
Post DRM Removal x Tail (Below 100k)	-0.683***		
	(0.062)		
Post DRM Removal x Middle (50k-500k)		-0.343***	
		(0.052)	
Post DRM Removal x Tail (below 50k)		-0.601***	
		(0.050)	
Post DRM Removal x Middle (25k-500k)			-0.397***
			(0.051)
Post DRM Removal x Tail (Below 25k)			-0.599***
			(0.051)
Observations	219,860	219,860	219,860
R-squared	0.409	0.408	0.407
Number of albums	3,307	3,307	3,307

Notes: This table presents the same results from Table 5 using triple interactions. In each column, the omitted category is the top part of the distribution, defined as an album selling above 1 million copies for Column (1) and more than 500,000 copies for Columns (2) and (3). Column (1) defines the tail as selling fewer than 100,000 copies, Column (2) is fewer than 50,000 copies, and Column (3) is fewer than 25,000 copies. The coefficient on the triple interaction that includes the tail is positive and significant in all three specifications, which suggests a statistically significant difference between tail albums and top-selling albums after DRM removal. All specifications include album and month-year fixed effects and a polynomial time trend of degree six.
Robust standard errors are clustered by album.
** p<0.01, ** p<0.05, * p<0.1

Table A2: The Impact of DRM Removal on the Sales Distribution (CEM sample)

Regression model: OLS	(1) Above 1M	(2) Above 500k	(3) 100k-1M	(4) 50k-500k	(5) Below 100k	(6) Below 50k	(7) Below 25k
EMI x Post DRM Removal	-0.106	-0.0142	-0.0580	-0.0277	0.209***	0.241***	0.233***
	(0.168)	(0.133)	(0.0731)	(0.0649)	(0.0604)	(0.0706)	(0.0706)
Observations	12,679	24,278	47,718	51,984	75,305	59,440	46,473
R-squared	0.370	0.338	0.270	0.221	0.138	0.140	0.153
Number of albums	202	388	767	842	1,253	992	775

Notes: This table presents the same results from Table 5 using the matched CEM sample. All specifications include album and month-year fixed effects and a polynomial time trend of degree six.
Robust standard errors are clustered by album.
** p<0.01, ** p<0.05, * p<0.1

Table A3: The Impact of DRM Removal on the Sales Distribution (based on ranks)

Regression model: OLS	(1) Top 200	(2) Top 300	(3) Top 400	(4) Middle (201-1000)	(5) Middle (301-1000)	(6) Middle (401-1000)	(7) Tail (>1000)	(8) Tail (>1500)	(9) Tail (>2000)
EMI x Post DRM Removal	-0.207	-0.025	-0.075	0.030	0.022	0.041	0.166***	0.185***	0.202***
	(0.205)	(0.171)	(0.142)	(0.109)	(0.117)	(0.129)	(0.043)	(0.047)	(0.051)
Observations	11,603	16,430	21,353	38,783	33,956	29,033	152,704	129,945	107,558
R-squared	0.406	0.377	0.366	0.294	0.293	0.286	0.171	0.159	0.152
Number of albums	183	260	338	616	539	461	2,508	2,142	1,777

Notes: This table presents the same results from Table 5 using a different definition of the sales distribution. I define different parts of the sales distribution by a rank ordering of albums based on the total sales in the first three years after album release. For example, the top 200 albums refer to the 200 albums with the highest three-year total sales after release. All specifications include album and month-year fixed effects and a polynomial time trend of degree six.
Robust standard errors are clustered by album.
** p<0.01, ** p<0.05, * p<0.1

Table A4: Technological Device Ownership Among Young and Old Americans

	18-29 years of age % Yes, have	65+ years of age % Yes, have
Smartphone	88	25
Laptop computer	79	41
iPod or MP3 player	63	16
Internet streaming device	62	15
Desktop computer	41	58
Tablet	34	25
E-reader	25	20

Source: GALLUP (2013)