
Laurina Zhang
Ivey Business School, Western University, London, Ontario N6G 0N1, Canada, lzhang@ivey.uwo.ca

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 (40%) but does not benefit top-selling albums. These results suggest that reducing search costs facilitates the discovery of niche products.

Keywords: copyright; long tail; technology; intellectual property; digitization

History: Received October 17, 2014; accepted May 18, 2016, by Lee Fleming, entrepreneurship and innovation.
Published online in Articles in Advance November 11, 2016.

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 et al. 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 10 years—from $14.6 billion in 1999 to $7.7 billion in 2009—and attributes the decline to piracy (Recording Industry Association of America 2010).

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 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

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 (Borland 2005). The resulting decision set the precedent such that producers of technology who promote the ease of copyright infringement can 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.
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 versus 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 this 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 dropped DRM from its catalogue in 2007, while the remaining major labels did 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 5,864 albums from 634 artists and is, to my knowledge, the longest and broadest panel constructed to describe music sales. The data include album-month-level data on the number of albums sold through offline (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 pretrends in the sales of EMI albums before DRM removal. Third, my estimations of the impact of DRM removal across different parts of the sales distribution rely only on variation within EMI’s catalogue of albums (rather than variation across labels). 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 the heterogeneous impact of DRM removal on different parts of the sales distribution.

My estimates suggest that dropping 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 (up to 40%) but does not benefit top-selling albums.

While a potential concern is that the increase in sales of long-tail albums may be due to the higher value of DRM-free music, I do not find evidence that the increase can be 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 sales of older albums. I find that long-tail albums of newer vintages experience a comparable increase in sales relative to long-tail albums of older vintages. Furthermore, long-tail 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 lowered search costs. As such, my results are consistent with theory that shows lowering search costs can facilitate the 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 2009, Danaher et al. 2014), 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 versus 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 (Aguirre 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 flows2 (e.g., Agrawal and Henderson 2002, Jaffe et al. 1993), cumulative innovation (Murray and Stern 2007, Williams 2013, Galasso and Schankerman 2015), and price (Li et al. 2013).

The rest of this 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 and 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 underincentivize 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. Indeed, the International Federation of the Phonographic Industry (2010) states that “music is an investment-intensive business… Very few sectors have a comparable proportion of sales to R&D [research and development] investment to the music industry.” The RIAA also states that “all these anti-piracy efforts are to protect the ability of the recording industry to invest in new bands and new music” (Associated Press 2005, quoted in).

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.3 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.

A growing body of literature has focused on the extent to which file sharing has displaced album sales. The theoretical literature on piracy and sharing has shown that piracy is not definitively bad for firms. For instance, while piracy can create an illegal source of competition and have a negative impact on rights holders’ profits, 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 et al. 2006).4 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 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 et al. 2009), consumers also have more options to buy less music since albums are

3 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, audiotape recording, and videotape recording had always been objected to by copyright holders and content producers. See Scherer (2003).

4 See Belleflamme and Peitz (2011) for a review.
“unbundled” online (Bakos and Brynjolfsson 1999). Moreover, consumers’ modes of music discovery, previously dominated by local radio, now include global streaming sites, social sharing, and recommendation engines. 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. For example, whereas Walmart may only carry the top 3,000 music albums that have the broadest mainstream appeal because of limited shelf space and local demand, online retailers such as 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. The long-tail argument suggests that niche content, such as older catalogues, remote genres, and music from lesser-known artists, is 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 et al. 2006), home video sales (Elberse and Oberholzer-Gee 2007), and music consumption (Bhattacharjee et al. 2007, Dewan and Ramaprasad 2012). Missing from the debate, however, is the link between the sales distribution of digital content and firms’ private responses 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. As the cost of copying digital content becomes lower, music publishers 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 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, many firms in creative industries use DRM to address piracy concerns.

The four major record companies—EMI, Sony, Universal, and Warner—which control the distribution of more than 80% of the music market (International Federation of the Phonographic Industry 2005), first required that DRM systems be implemented in conjunction with the emergence of online music markets to protect their music from being copied illegally. 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 through which to purchase music for download from the major music record labels.

The technical restrictions imposed by DRM aim to inhibit the use and distribution of digital content that is not intended by the content provider, whether this is through large-scale peer-to-peer sharing or small-scale sharing through family and friends (Belleville and Peitz 2014). 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. For example, 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 et al. 2011). Thus, the effect of DRM removal on album sales—and, in particular, on

---

8 Some book publishers are currently in the process of removing DRM from their catalogues, while the film and video game industries are actively using DRM on their catalogues. See bradevanrosen (2010), MacManus (2012), and Tassi (2012).

9 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.

10 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.

11 Sinha et al. (2010) describe that while DRM has been widely viewed by the music industry as an effective strategy for reducing digital piracy, its restrictions prevent consumers from gaining information about product fit and sharing files with friends. They formalize this in a model to show that varying the level of DRM conditions, such as ability to share files with friends, has a
the sales distribution—is ambiguous and an empirical question.

A major change occurred in April 2007, when EMI became the first major record company to remove DRM protection from 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 Fools’ joke. 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 any committee or board meetings within the industry.” This is largely because the major record companies have traditionally been staunch supporters of DRM technology. The concern that relaxing sharing restrictions will lead to higher levels of piracy was significant both before and well after EMI dropped DRM. 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 chief executive officer 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” (Gohring 2007, quoted in). Even after EMI made the announcement to drop DRM, commentators did not expect the other record companies to jump on the bandwagon, and many industry observers called EMI’s move “shortsighted” and “risky” in that “EMI haven’t tested it enough, so they don’t know what the market reaction is going to be . . . how will it effect piracy?” (according to an unnamed music executive, quoted in Angell 2007).

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 is 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 did 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.

4. Empirical Strategy and Identification

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 for 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 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 sales may be attributed to preremoval 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 three major record companies removing DRM in April 2009, I can employ a difference-in-differences strategy, where I compare the sales of similar albums with and without DRM before and after DRM removal. My main estimating equation is

\[
\log(AlbumSales + 1)_{it} = \alpha + \beta(EMI_i \times PostDRMRemoval_{it}) + \sum_{s=1}^{6} w_s t^s + X_{it} + \delta_i + \mu_t + \epsilon_{it},
\]

where I define \(AlbumSales\) as the number of digital tracks sold by album \(i\) in month \(t\) divided by 10, following industry standards. This is an appropriate measure of online sales because the majority of

---

12 Barney Wragg, phone interview by Laurina Zhang, April 30, 2014.
13 See, for example, Arthur (2007).
14 See, for example, Mnookin (2007).
15 See also Hart (2007), Kingsley-Hughes (2007), and Ulanoff (2008).
16 See Peoples (2011) for details.
online music sales are from digital downloads. I select online sales as the main dependent variable because DRM directly placed restrictions on how digital music was shared, but I also examine its impact on offline sales in some specifications. I log the dependent variable because album sales are skewed. Since the specification relates log of sales to the dummy variable, $EMI_i \times DRMRemoval_i$, I compute the marginal effect as $e^{\beta Y_i} - 1$. $EMI_i$ is a dummy variable equal to 1 for albums released by EMI; $PostDRMRemoval_i$ 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, $\beta$ 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. 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 ($s_i$) and month-year fixed effects ($t_i$). 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 6 ($t^s$), where $t$ denotes time in months, to flexibly control for differences in sales resulting from album release dates. Robust standard errors are clustered by album to reduce the potential for overstating statistical significance as a result of serial correlation within albums (Bertrand et al. 2004).

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 outcome 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 include the monthly physical sales of the album as a control. This measure of offline album sales is album specific and time varying, and it allows the influence of offline sales and any correlates with offline sales (e.g., offline popularity, decade popularity) to vary over time. Third, and importantly, the estimation on changes in EMI’s sales distribution 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 was 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 with 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 is 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.

My data cover music sold between January 1992 and June 2012. This data set contains monthly data on...
of the number of physical albums sold through traditional outlets such as 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 R&B, and jazz and classical. The unit of analysis is the album-month.

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.

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 more than 30 labels. Further complicating matters, SoundScan does not report the major record company, but rather the label under which the album is released. Thus, I manually match 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 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.

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 and 2006.

Like the sale of most creative products, such as books and movies, music sales are extremely skewed, particularly for offline album (i.e., Walmart) sales. An average album sells approximately 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.

I define five mutually exclusive parts of the sales distribution based on the total number of albums sold in the first three years after album release: albums that have sold more than 1 million copies (designated as “platinum” by the RIAA), between 250,000 and 1 million copies, between 50,000 and 250,000 copies, between 10,000 and 50,000 copies, and below 10,000 copies. An average album in the top part of the sales distribution (above 1 million copies) sells about 662 times more than an average album in the below

20 I do not include non-major music in my sample because there lacks a clean experiment involving DRM for independent music. This is for two reasons. First, indie music has always been available DRM-free since 2003 on sites that focus on indie music such as eMusic, CDBaby, CD Baby, etc. (Cohen 2004). Second, major music released on online distributors such as iTunes must comply with the majors’ stance on DRM protection. Any indie content on iTunes is also subject to the same DRM restrictions for the sake of consistency and experience of iTunes (Lewis 2007, Eran 2007). Thus, while independent music that was on iTunes gradually became DRM-free after April 2007, it had always been available without DRM since 2003 through other channels. This makes it difficult to cleanly identify the impact of DRM on indie music sales through a quasi experiment. 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. Furthermore, I believe the albums I am capturing at the tail of the distribution are representative of long-tail content sales. For example, the average album released by majors sells approximately 34,000 copies in the first three years of release, while the average indie album sells approximately 5,000 copies in the first three years of release (The Musical Disconnect 2011). About 14% of my sample comprises albums that sell below 5,000 copies, and 40% of my sample comprises albums that sell below 34,000 copies.

21 Typically, major labels have different labels for different genres. For example, Sony has several labels under its Sony Music Nashville branch to oversee their country artists.


10K category. A median album in my sample sells approximately 60,000 copies within the first three years of release. Thus, albums that sell below 50,000 copies (sales categories 10K–50K and below 10K copies) form a reasonable definition of long-tail albums in my sample.\textsuperscript{24} About 10% of the sample sells more than 1 million copies in the first three years of release, whereas roughly 46% of the sample sell fewer than 50,000 copies during the first three years of release; that is, the number of copies sold by the top 10% is almost 25 times more than what is sold by almost half of the albums in the sample.

The key explanatory variables are EMI and Post–DRM Removal. I use the indicator EMI to identify albums released by EMI. In other words, I use EMI to distinguish between my “treated” and “control” albums. Post–DRM Removal 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.

### 6. Results

In the sections below, I start by estimating the impact of DRM removal on EMI’s album sales. I examine whether the main results are driven by endogenous timing by investigating the presence of pretrends. I also discuss the extent to which the estimate can be interpreted as an average treatment effect. Then I investigate whether dropping DRM facilitates product discovery in the long tail by estimating 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 album sales. All baseline specifications include album

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Summary Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Variable name</strong></td>
<td><strong>Description</strong></td>
</tr>
<tr>
<td><strong>Sales characteristic</strong></td>
<td></td>
</tr>
<tr>
<td>Total Sales</td>
<td>The number of albums sold per month, including offline and online albums</td>
</tr>
<tr>
<td>Offline Sales</td>
<td>The number of physical albums sold per month (i.e., retail channels)</td>
</tr>
<tr>
<td>Online Sales</td>
<td>The number of track album equivalents sold per month (calculated by dividing total tracks sold per month by 10)</td>
</tr>
<tr>
<td>Total Sales—First three years after release</td>
<td>The number of total albums sold in the first three years of release (albums released before 2004)</td>
</tr>
<tr>
<td>Above 1M</td>
<td>Equals 1 if the total albums sold in the first three years of release surpass 1 million copies</td>
</tr>
<tr>
<td>250K–1M</td>
<td>Equals 1 if the total albums sold in the first three years of release are between 250,000 and 1 million copies</td>
</tr>
<tr>
<td>50K–250K</td>
<td>Equals 1 if the total albums sold in the first three years of release are between 50,000 and 250,000 copies</td>
</tr>
<tr>
<td>10K–50K (long tail)</td>
<td>Equals 1 if the total albums sold in the first three years of release are between 10,000 and 50,000 copies</td>
</tr>
<tr>
<td>Below 10K (long tail)</td>
<td>Equals 1 if the total albums sold in the first three years of release are under 10,000 copies</td>
</tr>
<tr>
<td><strong>Album characteristic</strong></td>
<td></td>
</tr>
<tr>
<td>Post–DRM Removal</td>
<td>Equals 1 after April 2007 for all albums</td>
</tr>
<tr>
<td>EMI</td>
<td>The album’s label is EMI</td>
</tr>
<tr>
<td>SONY</td>
<td>The album’s label is Sony</td>
</tr>
<tr>
<td>WARNER</td>
<td>The album’s label is Warner</td>
</tr>
<tr>
<td>UNIV</td>
<td>The album’s label is Universal</td>
</tr>
</tbody>
</table>

Notes: The sample covers 5,864 albums from 634 artists. These albums cover all four major record labels (EMI, Sony, Warner, and Universal). I define five mutually exclusive parts of the sales distribution based on the number of albums sold in the first three years of release for albums released before 2004. The unit of analysis is the album-month.

---

\textsuperscript{24} As a point of reference, \textit{Billboard Magazine} estimated that the average number of copies sold of an album released by a major label was 11,253 per year (The Musical Disconnect 2011). This implies that an average album sells at least 34,000 copies in its first three years of release.
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.

Panel A focuses on the effect of DRM removal on EMI’s online sales. Column (1) implements my main specification in Equation (1). 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. EMI × Post–DRM Removal 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 approximately 13%.

Column (2) excludes holiday and compilation albums 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 approximately 8%. Column (3) includes a polynomial time trend of degree 6 to flexibly control for differences in sales resulting from album release date.

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 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

Panel B focuses on the effect of DRM on EMI’s offline sales (e.g., retail outlets). Interestingly, I find that DRM removal is associated with a 19% increase in offline sales (panel B, column (3)). This suggests that online sales did not rise because dropping DRM made offline sales more attractive.

²⁵ Note that EMI, is not separately included in the regression because I identify it through the album fixed effects.

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.”
sales relatively less attractive, but rather that EMI’s overall music sales increased after DRM removal.\textsuperscript{27}

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 $\beta$ as a treatment effect. If there is no pretrend, then one would be more convinced that the main results are not driven by endogenous timing. To investigate the presence of pretrends, I estimate the equation

$$
\log(\text{AlbumSales}_{it} + 1) = \alpha + \sum_{k=1}^{7} \beta_k (\text{EMI}_t \times \mu_k) + \sum_{s=1}^{6} w_s t^s + X_{it} + \delta_i + \mu_t + \epsilon_{it},
$$

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 degree 6. All estimates are computed relative to the period of DRM removal. The indicator variable for treatment status interacted with the period of DRM removal 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 album sales in that quarter. Two findings stand out. First, although the pre–DRM removal sales pattern suggests that the average quarterly difference in sales between EMI and non-EMI albums is above 0, 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 album sales are about 28% higher than non-EMI album sales in the quarter immediately following DRM removal, this effect increases to roughly 40% one year later.

Another concern when estimating the effect of DRM removal on album sales is the emergence of the Amazon online store in October 2007. 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 that sells its content rather than DRM removal. In Tables A.1 and A.2 in the appendix, I drop all observations after October 2007 and find similar results. This suggests that the boost in sales is not driven by the availability of DRM-free tracks on Amazon. I also experiment with placebo dates that correspond with other label-level events. I control for other label-level events at EMI and at other record companies and find similar results (see Tables A.3 and A.4). This suggests that the increase in sales is not driven by other label-level changes.

While these analyses alleviate many of the major concerns associated with interpreting the estimate as a causal effect, there are still limitations in interpreting the result as an average treatment effect because of unobserved differences between EMI and other record companies. In the next set of results, I examine the heterogeneous effect of DRM removal on different parts of the sales distribution. This set of results relies on variation within EMI’s catalogue of albums and thus mitigates concerns that unobserved label-specific factors are driving the difference between EMI’s top-selling and lower-selling albums.

6.3. The Long-Tail Effect and Product Discovery

DRM removal materially 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

\textsuperscript{27} One interpretation of the larger magnitude of the coefficient for offline sales compared with online sales is that online sales are closer substitutes to piracy, and some online sales are displaced by piracy. Thus, the effect on online sales may be even larger than the results would suggest. However, I caution overinterpreting this result as the coefficients for offline sales are not statistically different from the coefficients for online sales.
songs with valuations above the price, relaxing sharing restrictions may lead to lower online sales.\textsuperscript{28} On the other hand, relaxing sharing restrictions may increase online sales by facilitating product discovery.

The argument for product discovery is as follows. Music is an experience good, and consumers need to test these to value them. For instance, Peitz and Waelbroeck (2006) and Chellappa and Shivendu (2005) argue that digital copies and nonauthorized sharing (whether it is from family and friends or from the Internet) can be useful in providing consumers information on characteristics of the music that they otherwise lack. Without the availability of digital copies, consumers have to make an uninformed choice and thus choose at random between the available products.

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. In particular, music from popular artists is often played on the radio and is also more readily uploaded and downloaded on file sharing networks (Liebowitz 2008). Conversely, unpopular artists are those artists whose music is known by a smaller fraction of the population. Thus, information on unpopular artists is harder to find.

DRM removal relaxes sharing restrictions, which means it is easier to distribute music through large-scale peer-to-peer sharing and small-scale sharing through family and friends. While most observers agree that DRM is unlikely to deter serious pirates or technically proficient users from uploading music to file sharing networks, it does raise the cost of casual sharing of music across family and friends. For instance, sharing playlists with friends is a lot more difficult with DRM.\textsuperscript{29} Some commentators note that the person-to-person curation of content is the new automation-assisted word of mouth, and DRM impedes this form of sharing.\textsuperscript{30}

The extent to which DRM removal impacts music sales is likely to be different for popular and unpopular artists.\textsuperscript{31} 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 obtained a precise signal of the music beforehand (e.g., through radio or on file sharing networks). Thus, relaxing sharing restrictions is unlikely to increase the sales of music from popular artists. By contrast, DRM removal lowers consumers’ sampling and search costs for unpopular music in two ways. First, unpopular music is more likely to be found on file sharing networks after DRM removal because it is more easily shared. Second, the sharing of less popular items is more likely to be impacted by the person-to-person curation of content across family and friends. Indeed, the word-of-mouth literature (Dellarocas 2003, Godes and Mayzlin 2004) finds that consumers rely on word of mouth for riskier actions, such as niche products. Furthermore, sharing from family and friends allows consumers to obtain a precise signal of the products that fit their taste best (Gans 2012). Thus, 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 et al. (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 will buy in the absence of sharing technologies, nor the set of consumers that will always pirate regardless of search costs.

Next, I examine whether the removal of DRM disproportionately benefits albums at different parts of the sales distribution. I limit the sample to albums released before 2004 and run the estimation on the period after 2004. Figure 2 plots quarterly treatment effects on the long-tail albums for the periods preceding and following DRM removal. It shows that EMI’s long-tail albums experienced a 60-percentage-point increase in the quarters after DRM removal relative to non-EMI albums.

Figure 3 plots the marginal treatment effects of DRM removal at different parts of the sales distribution using the estimates in Table 3 along with 95\% confidence intervals. Several findings stand out: First, I find that the marginal impact of DRM removal on the unpopular song. This is unlikely to alter your purchasing decision of the popular song as you have likely heard it already (through radio or file sharing) but increases the likelihood that you will purchase other music by that unpopular artist.
the top-selling albums (above 1M and 250K–1M) is close to zero and statistically insignificant, which suggests 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 platinum-certified albums is quite wide—between −0.38 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 (50K–250K) is small and statistically insignificant.

By contrast, Figure 3 suggests the size of the treatment effects for EMI’s long-tail albums is positive and significant. Albums in the 10K–50K and below 10K sales categories experience a 19%–39% bump in sales after DRM removal. Given that the album volume in these sales categories sells, on average, 24,863 and 4,130 copies, this suggests that DRM removal boosts sales in the long tail by about 1,611 to 4,723 copies, respectively. The 95% confidence intervals surrounding these estimates are also much narrower compared with the confidence interval for the estimate of top-selling albums. Furthermore, the increase in the tail of the distribution does not appear to be caused by substituting away from offline album consumption because of changes in the relative value of online music that is now DRM-free. Offline sales of albums at the tail of the distribution also experience a disproportionate increase in sales, consistent with discovery (Table 3, panel B).

Note that my results do not imply that sales from discovery offset 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 as a result of DRM removal.

These results are consistent with theory that suggests lowering search costs can increase the sale of niche products (Bar-Issac et al. 2010, Yang 2013) and sharing can facilitate product discovery (Peitz and Waelbroeck 2006, Gans 2012). Taking the point estimates of my results, some simple back-of-the-envelope calculations reveal that EMI sells an additional 99,111 copies after DRM removal. Given that the price stays relatively constant during this period and assuming each album costs, on average, $10, dropping DRM boosted EMI’s revenues by almost $1 million.

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

$\text{Note that my results do not imply that sales from discovery offset 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 as a result of DRM removal. These results are consistent with theory that suggests lowering search costs can increase the sale of niche products (Bar-Issac et al. 2010, Yang 2013) and sharing can facilitate product discovery (Peitz and Waelbroeck 2006, Gans 2012). Taking the point estimates of my results, some simple back-of-the-envelope calculations reveal that EMI sells an additional 99,111 copies after DRM removal. Given that the price stays relatively constant during this period and assuming each album costs, on average, $10, dropping DRM boosted EMI’s revenues by almost $1 million.}$
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 as a result of relaxed sharing restrictions, but rather because they are more likely to consume long-tail music 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 music 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 (see Table 4). Consider an EMI artist, such as 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

<table>
<thead>
<tr>
<th>Table 3</th>
<th>The Impact of DRM Removal on Different Parts of the Sales Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression model: OLS</td>
<td>(1) Above 1M</td>
</tr>
<tr>
<td><strong>EMI × Post–DRM Removal</strong></td>
<td><strong>A—DV: log(Online Album Sales)</strong></td>
</tr>
<tr>
<td>Observations</td>
<td>-0.0865</td>
</tr>
<tr>
<td>(0.152)</td>
<td>(0.107)</td>
</tr>
<tr>
<td>Number of albums</td>
<td>346</td>
</tr>
</tbody>
</table>

**Notes:** Panel A examines the impact of DRM removal on online album sales, and panel B examines the impact on offline sales. Column (1) restricts the sample to top-selling albums that have sold over 1 million copies. Columns (2) and (3) examine the middle part of the distribution: albums that have sold between 250,000 and 1 million copies and albums that have sold between 50,000 and 250,000 copies, respectively. Columns (4) and (5) examine the tail of the distribution: albums with sales between 10,000 and 50,000 copies, respectively. All specifications include album and month-year fixed effects. A polynomial time trend of degree 6 is included. Compilations and holiday albums are excluded. Robust standard errors are clustered by album. DV, dependent variable; OLS, ordinary least squares.

<table>
<thead>
<tr>
<th>Table 4</th>
<th>The Impact of DRM Removal on Non-EMI Albums of EMI Artists Released Before DRM Removal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression model: OLS</td>
<td>(1) Above 1M</td>
</tr>
<tr>
<td><strong>EMI_{artist} × Post–DRM Removal</strong></td>
<td><strong>DV: log(Online Album Sales)</strong></td>
</tr>
<tr>
<td>Observations</td>
<td>19,010</td>
</tr>
<tr>
<td>Number of albums</td>
<td>346</td>
</tr>
</tbody>
</table>

**Notes:** Non-EMI_{artist} 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_{artist} 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 DRM removal disproportionately increases the sale of non-EMI albums of EMI artists in the tail of the sales distribution. All specifications include album and month-year fixed effects and a polynomial time trend of degree 6. Robust standard errors are clustered by album. DV, dependent variable; OLS, ordinary least squares.

`∗p < 0.1; ∗∗p < 0.05; ∗∗∗p < 0.01.`
Figure 4  (Color online) The Impact of DRM Removal on Albums of Different Vintages

![Graph](image)

Notes. Each dot represents the magnitude of the treatment effect for old (released before 1992) and new (released after 1992) albums at each part of the sales distribution (estimates from Table 5), along with upper and lower bounds for 95% confidence intervals.

long tail but does not benefit their top-selling non-EMI albums. Table 4 shows the impact of DRM removal for platinum-certified albums (above 1M) is close to zero with a wide confidence interval, whereas the impact is much larger (16%–40%) and more precise at the long tail. 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 and Genres

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 heterogeneous impact of DRM removal on different types of albums in the long tail that are most likely to benefit from product discovery.

Figure 4 plots the magnitude of the treatment effects of DRM removal on albums of different vintages—albums released before (old) and after (new) 1992. Table 5 reports the corresponding regression results. 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 with newer albums.

Figure 4 shows that the overall increase in sales for older albums is approximately 23%, while the overall impact on newer albums is small and statistically insignificant. In addition, both figures show a disproportionate increase to long-tail albums and a statistically insignificant effect on albums in the top or middle part of the sales distribution. Specifically, long-tail albums of older vintages experience an increase of 24%–40%, and long-tail albums of newer vintages experience an increase in sales of almost 40%. However, the magnitude of increase for old albums is not statistically different from that for new albums in the below 10K category. Taken together, this suggests that while the overall impact on sales is larger for older albums compared with newer albums, product discovery is more likely to disproportionately benefit less popular albums at the tail of the distribution regardless of vintage.

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. Figure 5 plots the treatment effects using the estimates from Table 6. Anecdotal evidence suggests that hip-hop and R&B are the most pirated genres, while jazz and classical are the least pirated genres. 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 and 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 44%. Moreover, the estimated difference between hip-hop and jazz at the tail of the distribution is also statistically significant.

One alternative explanation may be that there is heterogeneity in unobserved user distaste for the inconvenience of DRM, which is correlated with unobserved heterogeneous demand for music in a particular segment. For instance, if the average jazz and classical buyer owns a large number of devices as a result of 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

34 In unreported regressions, I also examine whether the estimated changes are attributed to differences in quality at different parts of the distribution. I collect Amazon album ratings for a random stratified sample of albums in my sample and find that EMI and non-EMI albums receive very similar ratings at the same parts of the sales distribution. I do not find evidence that higher-rated albums benefit more from discovery.

35 Results are robust to other age cutoffs.

36 Because of the smaller subsamples of hip-hop and jazz, I consolidate several of the sales categories listed in Table 3 into three mutually exclusive categories, although results are consistent if I use the same categories as above or other cutoffs.

37 According to Stieben (2011). This is also confirmed from descriptive evidence in Oberholzer-Gee and Strumpf (2007) and from EMI’s surveys (Danaher et al. 2014).
One plausible explanation is that hip-hop consumers have better information about music in their genre than jazz and classical consumers and, as such, are not differentially impacted by relaxed sharing restrictions. In particular, hip-hop consumers are more aware of long-tail items in their genre than jazz and classical consumers. This is consistent with descriptive evidence that shows the younger demographic file share more intensely than the older demographic and that a larger fraction of hip-hop consumers are composed of younger consumers compared with jazz and classical consumers. This is consistent with descriptive evidence that shows the younger demographic file share more intensely than the older demographic and that a larger fraction of hip-hop consumers are composed of younger consumers compared with jazz and classical consumers.

It is also unclear whether hip-hop and rap listeners own fewer devices compared with jazz and classical consumers. Survey evidence shows that the average age of rap and hip-hop listeners is younger than the average age of jazz and classical listeners (Radio & Television Business Report 2012, Jazz Audiences Initiative 2011), and younger Americans are more likely to own more devices (Dugan 2013). 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 with hip-hop and R&B consumers.

Notes. This table examines the impact of DRM removal on old (released before 1992) and new albums (released after 1992). All specifications include album and month-year fixed effects and a polynomial time trend of degree 6. Robust standard errors are clustered by album.

**p < 0.05; ***p < 0.01.

Figure 5 (Color online) The Impact of DRM Removal on Albums of Different Genres

Notes. Each dot represents the magnitude of the treatment effect for hip-hop and jazz albums at each part of the sales distribution (estimates from Table 6), along with upper and lower bounds for 95% confidence intervals.
Jazz Audiences Initiative 2011). Thus, we can infer descriptively that a larger fraction of hip-hop listeners file share and may have different modes of discovery compared with jazz and classical listeners. For example, survey evidence suggests that jazz listeners learn about new music from radio and recommendations from family and friends (Jazz Audiences Initiative 2011) and, as such, may be more impacted by relaxed sharing restrictions. Taken together, this may suggest that certain demographics benefit more from discovery as a result of 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 lower-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, particularly 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. By contrast, DRM removal facilitates sharing 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. After all, sharing from the right people (i.e., friends and family) prioritizes information and facilitates better matches (Gans 2012).

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 et al. 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, video games) that highlights this tension. DRM is a unique anti-piracy measure because it is a strategy that is implemented by firms rather than by IP policy and law enforcement. 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. 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 framework. 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., and do small harm to 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 are 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 raised consumer welfare without reducing firm revenue. My results indicate that relative to a regime of DRM, EMI’s

40 Two-thirds of hip-hop consumers are between the ages of 18 and 34 (Radio & Television Business Report 2012), whereas only 17% of jazz and classical listeners are under the age of 45 (Jazz Audiences Initiative 2011).

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

42 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.

43 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.

44 See McLeod and DiCola (2011) for a discussion on how copyright may constrain musicians building on another’s prior works.
music sales increased on net and disproportionately for long-tail content.

My analysis is, of course, subject to limitations. My data focus on music released by major record companies. While the sample is representative of commercially relevant music, independent music plays an expanding role in the marketplace and arguably shapes a significant portion of the long tail. Furthermore, a growing number of streaming services (e.g., Pandora, Spotify), music videos (e.g., YouTube), and Internet radio facilitate new ways of product discovery and content curation. As such, the impact of DRM removal is likely to be less pronounced in the present day. 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 with 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 that of music, and there are likely fewer complementary ways to substitute for product consumption. For example, research shows that while file sharing has reduced offline sales, demand for live concerts has increased (Mortimer et al. 2012), and concert prices are sensitive to search cost reductions in secondary markets (Bennett et al. 2015). 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 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.

Acknowledgments

The author is grateful to her Ph.D. advisors for their insights and guidance over the years: Ajay Agrawal (chair), Avi Goldfarb, Joshua Gans, and Joanne Oxley. She also thanks Brett Danaher, April Franco, Alberto Galasso, Shane Greenstein, Doug Hyatt, Ariel Katz, Nicola Lacetera, and Matt Mitchell for helpful suggestions. The author gratefully acknowledges funding support from the National Bureau of Economic Research’s Economics of Digitization working group, the Martin Prosperity Institute, and the Centre for Innovation and Entrepreneurship.

Appendix

Table A.1 The Impact of DRM Removal on Online Album Sales (Observations Dropped After October 2007)

<table>
<thead>
<tr>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full sample</td>
<td>Until October 2007</td>
</tr>
<tr>
<td>EMI × Post-DRM Removal</td>
<td>0.126***</td>
</tr>
<tr>
<td>(0.0338)</td>
<td>(0.0308)</td>
</tr>
<tr>
<td>Observations</td>
<td>357,320</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.349</td>
</tr>
<tr>
<td>Number of albums</td>
<td>5,864</td>
</tr>
</tbody>
</table>

Notes: This table investigates whether Amazon disproportionately impacted DRM-free album sales by dropping observations after October 2007. All specifications include album and month-year fixed effects and a polynomial time trend of degree 6. Robust standard errors are clustered by album.

Table A.2 The Impact of DRM Removal on the Sales Distribution (Observations Dropped After October 2007)

<table>
<thead>
<tr>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Above 1M</td>
<td>250K–1M</td>
<td>50K–250K</td>
<td>10K–50K</td>
<td>Below 10K</td>
</tr>
<tr>
<td>EMI × Post-DRM Removal</td>
<td>–0.125</td>
<td>–0.0125</td>
<td>0.0511</td>
<td>0.257***</td>
</tr>
<tr>
<td>(0.125)</td>
<td>(0.0844)</td>
<td>(0.0639)</td>
<td>(0.0739)</td>
<td>(0.0642)</td>
</tr>
<tr>
<td>Observations</td>
<td>15,143</td>
<td>24,567</td>
<td>36,727</td>
<td>34,712</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.388</td>
<td>0.376</td>
<td>0.285</td>
<td>0.208</td>
</tr>
<tr>
<td>Number of albums</td>
<td>346</td>
<td>562</td>
<td>867</td>
<td>837</td>
</tr>
</tbody>
</table>

Notes: This table investigates whether Amazon disproportionately impacted DRM-free album sales by dropping observations after October 2007. All specifications include album and month-year fixed effects and a polynomial time trend of degree 6. Robust standard errors are clustered by album.

Table A.3 The Impact of DRM Removal on Sales (Controlling for Other Label-Level Events)

<table>
<thead>
<tr>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sony acquires 50% stake in BMG</td>
<td>Terra Firma acquires EMI</td>
<td>Sony acquires BMG and Terra Firma acquires EMI</td>
</tr>
<tr>
<td>EMI × Post-DRM Removal</td>
<td>0.088***</td>
<td>0.098***</td>
</tr>
<tr>
<td>(0.036)</td>
<td>(0.030)</td>
<td>(0.033)</td>
</tr>
<tr>
<td>Observations</td>
<td>250,664</td>
<td>250,664</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.378</td>
<td>0.377</td>
</tr>
</tbody>
</table>

Notes: This table investigates whether other label-level changes are driving the increase in sales. This table controls for two other label-level changes separately (columns (1) and (2)) and together (column (3)): Sony acquires Bertelsmann’s 50% stake in BMG, and Terra Firma acquires EMI. All specifications include album and month-year fixed effects and a polynomial time trend of degree 6. Robust standard errors are clustered by album.

Table A.4 The Impact of DRM Removal on the Sales Distribution (Controlling for Other Label-Level Events)

<table>
<thead>
<tr>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Above 1M</td>
<td>250K–1M</td>
<td>50K–250K</td>
<td>10K–50K</td>
<td>Below 10K</td>
</tr>
<tr>
<td>EMI × Post-DRM Removal</td>
<td>–0.159</td>
<td>–0.007</td>
<td>0.042</td>
<td>0.233***</td>
</tr>
<tr>
<td>(0.123)</td>
<td>(0.084)</td>
<td>(0.064)</td>
<td>(0.075)</td>
<td>(0.067)</td>
</tr>
<tr>
<td>Observations</td>
<td>21,618</td>
<td>35,168</td>
<td>53,041</td>
<td>50,445</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.359</td>
<td>0.289</td>
<td>0.231</td>
<td>0.152</td>
</tr>
</tbody>
</table>

Notes: This table investigates whether other label-level changes are driving the changes in the distribution of sales. This table controls for two other label-level changes: Sony acquires Bertelsmann’s 50% stake in BMG, and Terra Firma acquires EMI. All specifications include album and month-year fixed effects and a polynomial time trend of degree 6. Robust standard errors are clustered by album.

***p < 0.01.