It was Fifty Years Ago Today: Recording Copyright Term and the Supply of Music

Megan MacGarvie
Boston University, Questrom School of Business and NBER

John McKeon
Boston University, Questrom School of Business

Jeremy Watson
University of Minnesota, Carlson School of Management

July 20, 2021

It was Fifty Years Ago Today: Recording Copyright Term and the Supply of Music

Abstract

This paper examines the effect of the expiry of recording copyright on the supply of music - in the form of re-releases, availability in streaming platforms, and concert performances - by artists popular in the UK in the 1960s. We find that recording copyright expiry has different effects on a song's availability in different distribution channels. The lapsing of copyright leads to a large increase in the number of re-releases in physical formats, holding constant artist, age, and year fixed effects. However, when a song's original recording copyright expires, it becomes less likely to be performed in concert. Moreover, copyright status is not associated with differences in availability on the digital streaming platform Spotify. These results show that copyright has nuanced effects on availability, and can lead to different and even opposite effects on availability of a product across different distribution channels. They also show that within the context of digital distribution, the impact of copyright on availability differs based on the business model of a platform.

Keywords: Intellectual Property, Copyright, Music Industry

JEL Codes: O34

Copyright is critical to the strategies of firms that sell creative products such as music. Yet much of the empirical literature on the effects of copyright has studied content distribution channels from the pre-digital era, typically focusing on a single distribution channel.[1] Because technological change has led to a rapid and dramatic evolution of business models in copyright-intensive industries, more empirical research is needed to inform and adapt copyright regimes and firm strategy (Greenstein et al., 2013). This is particularly true in the music industry, where consumers access music through several distinct channels with disparate vertical structures.[2]

This paper examines the nuanced effects of recording copyright on the availability of music, using a unique data set that covers the musical releases, live performances, and digital availability of musicians that were popular in the UK in the 1960s. Within this context we examine the effect of sound recording copyright expiration on the availability of affected recordings. In addition, because artists can use live performances to promote album sales, we assess whether copyrighted songs are more likely to be performed in concert than songs of the same age for which recording copyrights have expired.

We find that copyright has different, and even opposing, effects on the availability of popular music dependent upon the nature of the distribution channel. Several of our findings depart from prior estimates of the effects of copyright, studies which do not capture availability through alternative channels more relevant in the case of music, and do not address the role of digital streaming platforms.[3] In contrast to most prior research, we are able to identify the impact of a change in copyright status during an artist's lifetime (and can therefore observe changes in the supply of performances). We also exploit the extension of copyright terms in 2013 to empirically disentangle copyright status from year and age effects, correcting for potential bias due to 50th anniversary effects.

We make several contributions. First, examining data on re-releases of songs distributed in retail channels, our results suggest a substantial increase, of approximately 160-340%, in the number of UK re-releases of songs once recording copyright expires, relative to songs of the same age and approximate vintage remaining under copyright protection.

This result is consistent with previous findings for books, which are also distributed using a retail model, and serves to replicate these previous studies in a new music industry context. But because recorded music differs from books in the number of distinct copyrights that must be cleared before re-issuing a song, these results show that the stimulating effects of public domain status on

[1]Prior research on books demonstrates that copyright expiry leads to an increase in the availability of books in print (Heald, 2014a; Reimers, 2018).

[2]The movie and music industries, among other core copyright industries (WIPO, 2015), generate value through retail sales, live performances (which now generate a large share of total revenue for the industry), and bundled digital streaming (Waterman, 2009; Mortimer et al., 2012). The industry is in transition from sales of primarily physical (CD and LP) albums to distribution through Digital Streaming Platforms (DSPs) which have stimulated revenues for the industry as a whole after years of decline.

[3]While demand for e-books has grown considerably, digitization and a low incidence of piracy have not disrupted the book industry's longstanding "retail" model of distribution in which physical books and digital download e-books are sold directly to consumers through retailers, while the predominant e-book retailer Amazon remains relatively open on the supply side.

availability are observed even when some copyrights (e.g. composition rights) remain in force.

However, results on live performances challenge prior findings on copyright and availability. Live performances play an important role in promoting album sales and streaming demand. Exploiting the fact that we observe variation in copyright status during the lifespan of artists in our sample, we show that public domain songs by a given artist are less likely to be performed in concert than copyrighted songs by that artist. This finding is consistent with a model in which artists choose to perform (and thus promote) songs that bring the highest sales royalties. In contrast to prior research which has suggested that extending copyright does not encourage the restoration, maintenance and distribution of pre-existing works (Buccafusco and Heald, 2013), this suggests that the negative supply effects of the extension on re-releases may be somewhat counteracted by a positive supply response in live performances.

Our findings on digital platform availability further depart from prior results on copyright and availability. We show that the typical DSP distribution model, with its comprehensive licensing of large catalogs of rights, negates the effects of copyright expiry for any individual song. In contrast to our finding that copyright expiry leads to greater availability of songs distributed via a retail model of physical copies, we find that popular songs with recording copyrights in the public domain are no more likely to be available on Spotify than songs whose copyrights have yet to expire. To be precise, we show that a song whose recording copyright has expired in the UK but not in the US is no more likely to be available in the UK than a song protected by copyright in both locations. However, it is the predominant streaming platform strategy rather than digital distribution per se that explains this difference, because estimates based on the availability of permanent digital downloads from Amazon (a "retail" style platform) show similar effects around copyright expiry to what is observed for physical releases. In addition, there is suggestive but not dispositive evidence that copyright could still limit the availability of the most obscure artists on digital platforms.

Taken together, these results show that copyright has nuanced effects on availability, and can lead to different and even opposite effects on availability of a product across different distribution channels. They also demonstrate that within the context of digital distribution, the impact of copyright on availability differs based on the business model of a platform. As the revenue shares, and relative importance, of dissimilar distribution channels shift in an industry in transition, it is important to understand how the impact of intellectual property rights on product availability may change as new channels rise in prominence.

1 Prior Literature

Researchers in the economics (Reimers, 2018; Li et al., 2018), information systems (Danaher et al., 2017; Aguiar et al., 2018), and legal literatures (Heald, 2008a; DiCola, 2013; Heald, 2014a) have investigated copyright's market restrictions in terms of product access, prices, and creator incentives. With a few exceptions, the latter aspect has received relatively little empirical research. MacGarvie and Moser (2015), for example, study publisher payments to authors around the British copyright term extension of 1814 and find that payments increased substantially after the extension, particularly for superstar novelists. Giorcelli and Moser (2016) show that Italian states that adopted copyright laws, as a result of annexation by Napoleon, saw a five-fold increase in the creation of historically significant operas and a ten-fold increase in the creation of operas still available in recordings today.

Product access includes consumers' access to copyrightable information goods as well as supply-side access – that is – the re-use of copyrighted material in the production of new information. A growing literature considers copyrights effects on supply-side access and reuse (Heald, 2008a, 2014a,b; Nagaraj, 2018; Biasi and Moser, 2018; Watson, 2018). Empirically focused papers in this stream consistently find that copyright protection decreases re-use across contexts, such as scientific works (Biasi and Moser, 2018), music (Watson, 2018), and Wikipedia (Nagaraj, 2018).

Consumers predominantly access music through retail sales (Rob and Waldfogel, 2006; Oberholzer-Gee and Strumpf, 2007; Hendricks and Sorensen, 2009), live performances (Mortimer et al., 2012; Cho et al., 2017), and digital platforms (Danaher et al., 2014; Aguiar and Waldfogel, 2018). In this work, we consider how the copyright status of a song affects consumer access through all three of these channels. Early research in the legal literature (Heald, 2008a,b, 2014a) demonstrates that copyright status is correlated with availability, but inference from these observational studies is limited by the absence of causal empirical design. Building on these works, Reimers (2018) utilizes a regression discontinuity approach and finds that lapse into the public domain results in an average of 26.5 additional book editions.

However, because the data used in these studies is drawn from the book industry, they do not study copyright's effect on availability aside from a retail (or bookstore) model of selling physical copies, or the similar model of permanent digital download sales through an open retail platform (or online marketplace, e.g. Luca (2017)). Alternative distribution strategies stemming from product differentiation, segmentation, and bundling are common in the copyright industries (Smith and Telang, 2009; Shiller and Waldfogel, 2011). Furthermore, while some content platforms exhibit a great deal of openness towards content providers on the supply side (e.g., YouTube, Steam), many of the dominant platforms such as Spotify and Netflix exert greater control. In this paper, we show that copyright status interacts with the nature of the distribution channel to influence the availability of music. With the shift to digital distribution of music, we show that copyright status may or may not influence availability, depending on the design of the platform on which music is distributed. This contributes to the growing literature on platforms, pointing out that digital distribution per se is less influential than the specific business models that digital distribution enables, and that the effects of platforms depend on the design choices made by the architects of platforms (Edelman et al., 2017).

2 Conceptual Framework and Underlying Economics of Copyright and Music Availability

In UK copyright law, a piece of recorded music may be protected by three separate copyrights. The first is the *musical composition*, or the pattern of notes, underlying the song. The rights to the musical composition are typically owned by the composer. The second aspect of copyright is the right to the lyrics to the song, which are treated as a literary work and typically owned by the lyricist. The third aspect is the *sound recording* which is the right to a specific fixation or recording of a song. Performers typically enter into contracts that assign recording copyrights to a record label in exchange for a royalty.[4] The sound recording can be thought of as the specific way that the song is performed and recorded. There can be multiple different sound recordings copyrighted separately for the same music composition performed in different ways or by different artists.[5]

Copyright in sound recordings (or "records, perforated rolls, and other contrivances by means of which sound may be mechanically reproduced") was established in the UK by the Copyright Act of 1911, which limited the duration of rights to fifty years (Copyright Act, 16/12/1911, Article 19, section 1). In 2013 the UK updated the Copyright, Designs and Patents Act 1988, following Directive 2011/77/EU, to extend the duration of sound recording copyright from 50 years to 70 years from the year of first publication. According to a post-implementation review of the legislation by the UK Intellectual Property Office, "The primary objective of the legislation was to enhance the welfare of performers (artists) and record labels, ensuring they receive appropriate rewards for their effort throughout their lives for their sound recordings and performances"(UK Intellectual Property Office (2018)).[6,7] This is still a much shorter term than the United States', which protects the copyright of sound recordings for the artist's life plus 70 years. The law extends protection on songs first published in 1963 or later, and EU member states were required to comply with the Directive by November 1, 2013.[8] Songs originally released before 1963 are unaffected by the extension and entered the public domain after 50 years.

Copyright affects different artists and songs in different ways. Artists that write, compose, and record their own songs will receive revenues from both the musical composition and sound recording

[4]Directive 2011/77/EU, section (9). Non-featured performers (i.e. session musicians who play in the background), who typically received lump-sum payments rather than a royalty, became entitled to receive royalties 50 years after the recording when the directive came into being.

[5]"Cover versions" are a common example of sound recordings that are distinct from the original version's recording copyright.

[6]This justification is focused on ex post compensation of artists rather than any motivation for stimulating creation of new work. Singer for The Searchers Mike Pender has recently said, "It is nice to still have the royalties. It means I can still take my wife to a nice restaurant!" `https://www.dorsetecho.co.uk/leisure/17519726.founding-member-searchers-mike-pender-reveals-spends-royalty-cheques`, accessed 5/24/2021.

[7]However, according to Theofilos (2013), "[m]ost artists who were young and just starting their careers were systematically forced by powerful record companies into signing deals that paid only low royalty rates and effectively forced those artists to relinquish all other rights to their music." Theofilos notes that Kretschmer (2011) finds that approximately 72% of the monetary benefits of term extension will go to record labels, with only 28% going to artists (and only 4% to artists facing an income gap). Stanley (2011) notes that the more popular artists would have renegotiated their contracts with labels and stood to gain substantially from the extension, particularly artists like Cliff Richard who did not write their own songs and therefore did not benefit from composition royalties.

[8]Article 2, section 1, Directive 2011/77/EU.

rights to their music, while artists that perform compositions written by others depend on the royalty stream from sound recording copyrights. Regardless of the nature of the benefits to the artist, a third party must pay for recording as well as composition rights when reissuing music. In what follows, we examine how copyright protection and its expiration affect the supply of music in terms of physical releases, availability on digital channels, and live performances in concert. These distribution channels represent the majority of current revenue streams for musicians.[9] We focus upon copyright's role in markets for existing products and do not consider copyright's incentives to create new products, a question outside the scope of this paper.

2.1 How Might Copyright Expiration Affect the Supply of Music in Physical Channels?

Following the prior literature on copyright and availability in books (Heald, 2014a; Reimers, 2018), we measure the availability in terms of the number of re-releases of a song. The decision to re-release a song is made by a label. Record labels, which control the recording copyrights of associated artists, may exclude others from using, exploiting, or distributing their songs for a statutory duration. Labels may exploit these copyrights for a monopoly profit for the duration of the copyright, or instead license the rights to entrants, with licensing fees set at a level such that the label is indifferent between options. We term these entrants "reissue" labels. Large established labels in the music industry possess complementary assets such as distribution networks, promotional capabilities, and a portfolio of copyrights that may be bundled together in a release. These complementary assets, combined with control of recording copyrights, ensure that major labels have an advantage over small reissue labels when releasing on-copyright songs. "Generic" reissue labels lack these assets and instead pursue a strategy of issuing lower-cost releases based on public domain works. As described by Synovitz (2003), regarding recordings that were public domain in Europe but rights-protected in the U.S.,

> "Without the requirement of paying royalties to the original record producers, firms like the U.K.-based Charley Records, Holland's Disky Communications, and France's EPM Remastering are able to sell public-domain CDs for as little as $2 each. In America, where the producers of the same recordings still exercise copyright control, music collectors routinely pay $15 to $25 per CD for the music."[10]

These entrants may distribute reproductions of the original recording, near-perfect substitutes for the original, at low marginal cost. Such reproductions may be sold either directly to the consumer by the entrant label or through retail outlets. Entry, and the associated increase in supply from physical recordings, should lead to lower average prices for consumers due to competition between close substitutes for a public domain recording. We observe only releases, not prices, and so do

[9] Because we lack data on synchronizations (3% of U.S. recorded music revenue) and terrestrial radio airplay, we do not examine the impact of copyright on these channels. Synchronizations are licenses for attaching music to other forms of media (e.g. TV, film, advertisements, etc.)

[10] Synovitz (2003) (https://www.rferl.org/a/1101848.html)

not estimate this latter price effect, which has been documented observationally elsewhere (Pollock et al., 2010).

According to the above, we should expect to see more reissues of a recording once it enters the public domain. However, some have argued that extending copyright terms may increase availability because it creates incentives for rights holders to invest in the restoration and public distribution of their works (Eldred v. Ashcroft, 537 U.S. 186, 207 (2003), cited by Heald 2014a).[11] For example, recorded music was distributed via successive standards of phonograph records and compact cassettes before the proliferation of digital audio CDs, DVDs, and current digital distribution over networks. Recordings must also be remastered to adapt to changing consumer tastes and listening technology (e.g., mono vs stereo recordings). Prior observational research has not found support for this idea (Buccafusco and Heald, 2013).

2.2 Promotion and Live Performances

Artists perform live concerts to generate income from ticket sales, as well as to promote sales of recordings (Mortimer et al., 2012). When choosing a set of songs to perform in concert, artists consider their private rewards for performing each song. These rewards are a mixture of monetary and non-monetary rewards (i.e., personal taste of the artist). Artists typically receive royalties from labels on each sale of a recording protected by copyright (see section 2), while recordings in the public domain generate no royalties. These artists have an incentive to prioritize the promotion of recordings still protected by copyright over those in the public domain, all else equal. This is clearly articulated by Cliff Richard, quoted in an article about the looming expiry of his recording copyrights:

> "It seems terribly wrong that 50 years on they lose everything from it...Sometimes I'm absolutely fed up with singing *Living Doll* but I have sung it constantly since 1959 because every time I sing it live it generates sales of the original record and royalties to me." (Cliff Richard quoted by Miller (2006))

The monetary rewards to performing a given song also depend on matching setlists to audiences tastes to ensure continued audience patronage. The expiration of a recording copyright could potentially affect this monetary reward as well, since labels may choose to promote and market only on-copyright songs over public domain songs, which may predispose audiences to prefer these promoted songs that are still under copyright.

It is worth noting that incentives for performance of back-catalog songs by older artists differ from the incentives faced by contemporary emerging artists. Some observers have suggested that album releases are primarily used to promote concerts, rather than vice versa, since in the era of internet piracy new artists derive most of their revenues from touring (Papies and van Heerde,

[11] One possible example that could be cited in support of this idea is the recent 50th anniversary reissue of the Beatles' *Sgt. Pepper's Lonely Hearts Club Band*, which included previously unreleased takes of all songs as well as remixes and 33 additional recordings from the original recording sessions.

2017).[12] This argument is less relevant, however, for artists that depend on a large back catalog of works, who have little or no new content to release, and whose audience primarily consumes music through legal channels rather than through piracy.

It is also important to note that the protection of a recording copyright does not limit an artist from performing a song publicly. Our analysis is not focused on performance rights which require permission to perform a song to be obtained from the holder of the composition copyright. Venues are responsible for these licenses, and typically acquire a blanket license administered by a national collection society (i.e., PRS for Music in the UK). Given the licensing context and the fact that these copyrights endure for 70 years (as described above), they are not relevant for the setlist decisions in our sample.

2.3 Copyright and Supply in Digital Channels

Streaming music royalties represent a source of growth for an industry that has faced declining revenues since the advent of peer-to-peer file sharing. Music distribution through streaming is concentrated among just a few large platforms, especially Spotify, Apple Music, Amazon, Pandora, and YouTube (Steele, 2018).

Since the value of a platform is greater to both buyers (listeners) and sellers (labels) when large numbers of buyers and sellers are present on the platform (Rochet and Tirole, 2003), a notable feature of DSPs is the availability of a near-complete catalog of music. Consumers pay a fixed monthly subscription, or view ads, in exchange for a bundle representing unlimited streaming of the entire catalog of music licensed to the DSP. Nicolaou (2018) quotes Matt Pincus, chief executive of music publisher SONGS who says that,

> "[t]he problem with audio streaming is you need to have 100 per cent of the content rights in the world for music, otherwise nobody buys your product."

Record labels have entered into catalog-encompassing licensing agreements with DSPs, such as Spotify, that enable this bundling strategy. DSPs such as Spotify do not search for and aggregate content themselves, but instead receive sound recordings delivered by major labels or intermediaries.[13] These labels and consolidators typically license entire catalogs to the DSP because "We have to be where the fans are" (Scott Borchetta, chief executive of Big Machine Label Group, quoted by Nicolaou 2018).[14] Presumably to help facilitate this comprehensive licensing approach, Spotify has given major record labels large advances on royalty payments and an 18% equity stake

[12] Papies and van Heerde (2017) cite the example of Prince, who released the album 20Ten for free to promote concert ticket sales.

[13] Small independent labels wishing to distribute content through digital platforms must contract with a consolidator such as CDBaby, which acts as an intermediary between labels and platforms. CDBaby delivers master recordings to Spotify and distributes payments to indie labels after taking a cut.

[14] Holdouts are exceedingly rare. Some superstar artists with distribution control, such as Prince, objected to widespread streaming of their catalog, though Prince's estate has since licensed the artist's catalog for streaming. Withholding new releases from DSPs has also been used in the past as a form of intertemporal price discrimination.

in the company (Cohen et al., 2015).[15]

Streaming platforms' usage of statistical learning to derive competitive advantage further encourages the universal availability of songs. Spotify derives revenue from advertising and subscriptions, but like other streaming platforms (e.g. Netflix as described by Smith and Telang (2016)), part of this strategy depends on improved recommendations to users and analytics services to labels and performers (Marr, 2017). The latter services depend on amassing troves of data across catalogs of songs and millions of customers. As a result, it is critical for Spotify to be able to offer comprehensive access to virtually all recorded music.

On the supply side of the platform, DSPs govern quality by limiting direct access to the three major labels and a set of large preferred distributors (Boudreau and Hagiu, 2009). This private regulation prevents excess variety from persisting on the platform, moderating consumer search costs on the demand side but excluding public-domain and reissue labels that thrived when CDs were distributed in physical channels. The platforms further limit the variations of the same song allowed on the platform to help consumers identify the song most likely to be searched.[16]

Thus, the decisions faced by DSPs contrast in several key ways with the model described above for physical channels, in which labels maximize profits by choosing the price at which to sell discrete quantities of albums, which vary in their cost of production according to copyright status. Spotify primarily derives revenue from advertisers and subscribers, and its users are attracted by near-universal access to music. Spotify maximizes an objective function in which the number of listeners using their product weighs very heavily, which creates strong incentives to make as much music available as possible, independent of copyright status.[17] Because the vast majority of recorded music remains copyright protected in developed countries, it would be infeasible for a general audience platform to favor lower-cost public domain songs over rights protected music.

Rights holders thus transfer bundles of on-copyright recordings with public domain recordings, with no space for a potential entrant label to enter with public domain recordings once the DSP has negotiated a catalog-wide license and transfer with the original rightsholding label. With this structure, potential entrants selling re-releases of public domain recordings could instead create their own DSP, but face high fixed costs along with stiff competition against incumbents in a network-effects driven market dominated by a few major platforms. In retail channels, public domain labels can provide value by distributing releases with unique combinations of public domain works, but the playlist and recommendation feature of DSPs, along with policies limiting excess variety, have

[15] Insight into the nature of licensing arrangements with major labels can be gained from a 2011 contract between Sony Music and Spotify which was leaked to the press, `https://www.theverge.com/2015/5/19/8621581/sony-music-spotify-contract`.

[16] These variations may be various remasters, mono vs stereo recordings, explicit vs clean recordings, radio versions vs album versions, and so on.

[17] Note that other music platforms pursue different strategies. Amazon music has both a streaming option and a retail (a-la-carte) model, as does Apple through its Apple Music streaming service and permanent digital downloads through the iTunes store. Non-interactive services, such as Pandora, may be less reliant on a complete catalog of music as the platform chooses songs for the listener. Although licensing sound recordings on non-interactive services follows a statutory rate in the U.S., Pandora at one point negotiated lower royalty rates for a network of independent labels (Merlin) in exchange for steering listens to these independent artists. See the written statement of Pandora Media, Docket No. 14-CRB-0001-WR (2016-2020)

disrupted the strategy of these small entrant labels. The blanket licensing of large catalogs of rights may thus negate the effects of copyright expiration for any individual song. This model stands in contrast to the distribution model for physical sales, in which even small reissue labels can, at relatively low cost, distribute a compilation CD of public domain songs.

However, despite the incentives for universal availability in the DSP model, it is not necessarily obvious that public-domain and copyrighted songs will be equally available. The long life of copyrights may still pose challenges for the digitization and distribution of information from past generations.[18] First, musicians face non-pecuniary incentives, and may resist new forms of digital distribution for personal reasons.[19] So long as copyright protection of a recording persists, any residual control rights of the original artist may complicate licensing and prevent distribution. Further complications arise from aging recording and distribution contracts between record labels and artists, which may have ambiguous terms regarding the control of digital distribution rights, and thus the recordings and rights covered by aging agreements may be omitted from catalog-wide licensing agreements. The original label may be defunct, and the contemporary rightsholder may be difficult to identify or locate, similar to the "orphan works" problem wherein content remains unexploited because rightsholders cannot be located by a diligent potential licensee (Varian, 2006).

While major labels may have less difficulty locating, interpreting, and renewing contracts with the most prominent artists of high-value songs, agreements made more than 50 years ago between small independent labels and more obscure artists may be harder to locate and interpret. If old licensing agreements must be renewed or renegotiated for digital distribution, transaction costs may overwhelm the expected profits for out-of-fashion music. These incomplete contracting issues have afflicted the digital availability of even more recent artists. In an interview with Bloomberg, Kevin Mercer of the group De La Soul spoke of the contractual issues surrounding the group's albums released in the late 1980s and 1990s,

> "Unfortunately, a lot of the earlier stuff we did on Tommy Boy [Records], from what we understand, a lot of the legal language that needed to be a part of the contracts between ourselves, the owners of the master and the publishing, I guess it didn't include the world of digital. It was almost specifically to vinyl, cassettes, CDs. So a lot of those contracts needed to be reworked."[20]

Such uncertainty around property rights is eliminated when a recording enters the public domain, and this may increase the likelihood of availability on the platform. If it is true that public domain status increases the likelihood of a song being available on Spotify due to reductions in uncertainty about the ownership of rights, we would expect this effect to be primarily observed among the relatively obscure artists for which licensing transaction costs may be high compared to the expected value of streaming. We test this hypothesis in the empirical analysis.

[18] For example, the litigation regarding Google Book Search in *Authors Guild, Inc. v. Google, Inc.*

[19] See, for example, https://www.billboard.com/articles/news/7549739/artists-streaming-services-holdout-not-available.

[20] https://www.bloomberg.com/news/videos/2015-04-09/de-la-soul-raises-400k-on-kickstarter

3 Data

3.1 Physical Release Data

In order to identify a set of songs affected by the copyright term extension, we collected all Top 20 UK Album charts from 1960 through the end of 1965 from Officialcharts.com. OfficialCharts provides a top 10 list of UK albums for the first 11 weeks of 1960, whereas for the remainder of 1960-1965 it provides weekly top 20 lists. These charts were then carefully hand-matched to the MusicBrainz database (musicbrainz.org) to link the artists in the OfficialCharts data to the unique artist identifiers in the MusicBrainz database. There are 138 artists from OfficialCharts that match to the MusicBrainz data, and 45 artists that did not appear in the database or had no relevant releases. Soundtrack albums appearing on OfficialCharts were excluded, including albums credited to "Original Soundtrack," "Original Cast Recordings," "Original Broadway Cast," etc.

Using the sample of artists gathered from OfficialCharts, we then collect all songs released by these artists in the MusicBrainz database. We collect: artist name, release name (e.g., the name of the album/EP), the country of release, the date of release, the song name (standardizing case and stripping accent marks), the release type (album/single/ep), whether the song is part of a re-release, and the year of original release.[21] A song is considered a re-release if there is an exact match for the artist and song title with a prior date in the database.

In the UK, sound recording copyright terms begin with the original release of a recording. Year of original release, for the songs in our sample, runs from 1923 to 1975.[22] We create a final dataset in which the unit of observation is at the song-year level, and the key dependent variable is the number of re-releases of that song i in the UK in year t.

Information on the record label is available for approximately 90% of the UK recordings in our dataset. We classify a release as a major label releases if the *label type* field on MusicBrainz is one of production, original production, imprint, or holding. Reissue Production labels are the second most common label type in our database. There are clear patterns of specialization by label and copyright status of songs. Major labels are dominated by a few large firms such as Columbia, EMI, Parlophon, and Virgin, and mostly issue copyrighted songs. Reissue labels (such as Real Gone Jazz/Real Gone Music or Not Now Music) represent only 20% of releases for copyright protected songs, but 43% of releases for songs in the public domain.

We exclude observations on songs re-released more than 54 years after the original release year, because this is the maximum age observed for songs released in 1963 in our dataset (since data ends in 2017). As seen in Table 1, the final dataset used for our regressions contains 922,182 observations

[21] To ensure accuracy, we also consider the year of original release from Discogs, and use the earlier of the two years. Further, we manually confirm original release years for any song in the random sample used for Amazon analysis (described below), performed after age 40, or re-released after age 40 by an artist with an error rate above 20% within the Amazon sample. Validation of original release years was done using sites such as Wikipedia, Secondhandsongs.com, and 45cat.com

[22] The releases before 1940 are by Louis Armstrong, Bing Crosby, Duke Ellington, Ella Fitzgerald, Judy Garland, Count Basie, Glenn Miller and Frank Sinatra.

on 18,516 recordings of songs by 138 artists. We observe songs every year after the original release year from 1960 to 2017, and the average age of an observation in the dataset is 26.6. The typical song-artist pair is observed 49.8 times in our final dataset.

3.2 Set List Data

We obtain data on songs performed in concert from www.setlist.fm, a wiki service on which users post lists of songs performed in concert. We queried this site's API for songs of artists by MusicBrainz ID for all the artists in our reissue database. We then matched song names listed in MusicBrainz to the performed songs listed on setlist.fm. We created a crosswalk of standardized names by parsing out extraneous characters and standardizing case to match songs between the two datasets. We exclude from the database any artists and songs that never appear on setlist.fm, because for these artists we cannot distinguish between songs that were never performed and songs that were not recorded by the creators of the setlist database.

Table 2 reports summary statistics on the set list data. We create one observation per year that the song could have been performed from 1960 to 2016 and exclude observations in which the song is at least 55 years old. We create a count variable containing the number of times the artist performed song i in year t. The average number of performances of a song in the UK in a given year is 0.244. The typical song is performed in 15.3% of potential song-years. We also restrict the data to performances in years before the artist's last active year, according to MusicBrainz. After these limitations we have 109,583 performance-year observations across 84 artists and 2,825 songs. We flag years in which the artist was on tour according to our set list data so that we are able to run analysis conditioned on touring as well.

3.3 Streaming Data

Artists in our MusicBrainz data were hand matched to Spotify's artist unique identifiers (URIs). Catalog information, including geographic availability, was then downloaded from Spotify's API for all of the artists in our MusicBrainz dataset, and song titles were matched between the two datasets (see the Data Appendix for details). The resulting match allows us to identify which songs are available on Spotify, by year of recording and by location of the Spotify user.[23] Our hypothesis is that pre-1963 songs will be more available in the UK relative to the US due to their public domain status in the UK.

In contrast to the release and set list datasets, which are panels, the Spotify data is a cross section reflecting availability on Spotify as of September 2017. Unlike re-release and set list data, we do not observe songs suddenly reaching the age-50 cutoff in this dataset, because it was collected as a cross section after the term extension directive was enacted. There are a total of 37,032 observations in

[23] Although it is probable that a small percentage of potential matches were missed in our matching procedure, it is unlikely that the missed matches are in any way related to copyright status, especially since we can observe a song's availability in the US as a baseline estimate.

the Spotify dataset (one observation on availability in the UK market and one on the US market for 18,516 songs). The variation in copyright status comes from the fact that recordings made before 1963 are in the public domain in the UK, while recordings made after that date are under copyright protection in the UK, and none of the recordings in the dataset are in the public domain in the US.

Summary statistics on Spotify availability for artists found in the Official Charts Company (OCC) top albums charts are found in Table 3. For robustness, we also collected information on more obscure singles, also found in Table 3.[24] Of the 18,516 songs in the main sample, 72.2% are available in the US, and 77.0% are available in the UK. Of the 7,027 songs in our sample recorded prior to 1962, 76.5% are available in the US market on Spotify, and 82.1% are available in the UK. For the 11,489 songs recorded between 1963 and 1975, 69.7% are available in the US and 73.9% are available in the UK. For the median artist in our sample, 85.3% of the artist's songs are available either in the US or in the UK.[25]

4 Estimation and Results

4.1 Empirical Model

In this section, we examine the effect of public domain status on the availability of songs. We begin by estimating the impact of songs lapsing into the public domain on the availability of physical releases and live performances, before estimating the effect of public domain status upon a song's availability on the dominant music streaming platform Spotify. Very generally, we estimate the following model where Y measures the availability of song i, $PublicDomain$ is an indicator for public domain status in the UK, β controls for year effects, γ for song-age effects, δ for artist effects, and ε captures idiosyncratic shocks. In an alternative specification, we will use ζ to control for individual song effects. It is important to note that our data on physical releases and live performances are panel data, with songs originally released in year t_0 and potentially re-released or performed live in a later year, t. The age of a song in year t is thus $A = t - t_0$.

$$Y = f(PublicDomain; \beta, \gamma, \delta, \varepsilon) \tag{1}$$

Before estimating the effect of copyright status on the availability of works, it is useful to consider the ideal experiment that would be used to study this question. In an ideal experiment, the econometrician would select a sample of copyrighted recordings from the population of recordings, and randomly partition this sample into a treatment and control group. While the control group remains undisturbed, the econometrician would nullify the recording copyrights of the treatment

[24] To assuage concerns that our Spotify dataset consists of only the most durable artists, we collected an additional sample of songs by more obscure artists for the purpose of robustness checks. These artists are drawn from the OCC lists of weekly top 50 singles charts between 1960 and 1965. These additional data represent a cross section as of July 2019.

[25] The artists with availability rates below the 5th percentile of 35% are Bert Weedon, Big Ben Banjo Band, Paddy Roberts, The Dave Clark Five, The George Shearing Quintet, and Wayne Fontana and the Mindbenders.

group. With this treatment condition, the availability of recordings could be compared between the copyrighted control group and the public domain treatment group in terms of physical reproductions, availability on digital streaming platforms, and live performances.

In contrast to the ideal experiment, we use observational data, in which recording copyrights end after a statutory duration. Previous research demonstrates a correlation between a song's popularity and song age (Waldfogel, 2012), which could confound an analysis of the effects of public domain status. While one might think that a regression discontinuity design with a break at age 50 would be sufficient to identify the effect of entering the public domain, such a design is complicated by the fact that the fiftieth anniversary is a common time for revisiting long-lasting cultural products.[26] We might therefore expect to see a flurry of commemorative re-releases or performances even in the absence of copyright term expiry. This flurry could lead to biased estimates of the effect of copyright, unless we control for age.

Although panel data allows the econometrician to account for the aforementioned age effects, it is also important to control for the shift in the music industry towards digital distribution. An analysis based purely on age would confound the fact that songs reaching age 50 in the years after 2010 are entering the public domain at the same time as demand for physical re-releases was declining relative to digital distribution. Without year controls, an analysis based purely on age will underestimate the impact of the public domain on availability.

Our analysis exploits three different sources of variation to identify the effect of copyright on availability. First, we estimate the change in availability when a song enters the public domain using only the period prior to the copyright term extension, in which all songs enter the public domain at age 50. Then, we extend to the period after the term extension and control for age and year effects. Finally, we estimate a triple difference regression in which US observations on a song serve as a control group for UK observations on that song.

Our preferred approach is the second, which controls for both age and year effects and relies upon the variation in copyright status created by the term extension. For recording copyrights published after November 1963, the law exogenously extended the term of recording copyrights from 50 years to 70 years. Recording copyrights published before November 1963 lapsed after 50 years, and were unaffected by the term extension. The copyright term extension thus allows us to separate the effect of copyright status from shifts in demand as songs grow older, as well as anniversary effects. Absent the term extension, copyright status in our sample would be perfectly predicted by recordings aging past 50 years. While all recordings in our panel below 50 years of age are protected under copyright, this extension provides a sample of counterfactual recordings that are greater than 50 years old and in the public domain, as well as recordings of the same age that are still under copyright protection, which allows us to estimate the counterfactual of availability under longer copyright terms.

With this research design in hand, we estimate the effect of copyright protection on the re-releases

[26] See, for example, the recent 50th anniversary re-release of the Beatles' *Sgt. Pepper* album. Another example is *Genius & Soul - The 50th Anniversary Collection*, a Ray Charles album released in 1997, containing songs from the 1940s to the 1990s.

or performances of a song, using the following model:

$$Y_{itjA} = \alpha_0 + \alpha_1 PublicDomain_{itjA} + \beta_t + \gamma_A + \delta_j + \varepsilon_{itjA} \quad (2)$$

In which the dependent variable is the count of re-releases or performances in the UK of song i by artist j in year t with song age A. Because the dependent variable is a count, we use Poisson pseudo-maximum likelihood regressions with multi-way fixed effects (Correia et al., 2019). $PublicDomain_{ijtA}$ is a binary variable equal to one if the song's recording copyright has expired in year t and equal to zero if it is still under copyright protection in year t.[27] β_t is a dummy for release year t, γ_A captures the fixed effect of song age, and δ_j is the artist fixed effect. In contrast to prior studies of copyright term extensions that performed before-after analyses, the 2013 term extension allows us to control for a full set of age and year effects.

In alternative specifications, we control for song fixed effects, ζ_i and year and age effects:

$$Y_{itjA} = \alpha_0 + \alpha_1 PublicDomain_{itjA} + \beta_t + \gamma_A + \zeta_i + \varepsilon_{itjA} \quad (3)$$

The song fixed effects control for all song-level characteristics including artist and year of original release.

The data on live performances and physical releases have a panel structure, with observations on performances and releases over time. In contrast, Spotify digital streaming availability is observed only in a cross-section of availability as of 2017.[28] Instead of comparing variation in availability before and after copyright expiry, we instead exploit variation in copyright status between the US and UK. We assess whether there are differences in availability between the US (where all songs in our sample are still protected by copyright) and the UK (where recordings made prior to 1963 are in the public domain) in our cross-country cross sectional data.

In order to measure the effect, we estimate the following model:

$$Y_{ijAk} = \alpha_0 + \alpha_1 PublicDomain_{ijAk} + \gamma_A + \theta_k + \delta_j + \varepsilon_{ijAk} \quad (4)$$

In which the dependent variable is equal to 1 if song i by artist j of age A (that is, originally released in year t_0) is available on Spotify in country k (either the UK or the US) in the sample year. $PublicDomain_{ijAk}$ is a dummy equal to 1 for UK observations on songs originally released before 1963, and 0 otherwise. Because we only observe these songs in a single year, we no longer control separately for age and year effects, and instead include in some specifications γ_A to control for age fixed effects as in previous specifications. δ_j again controls for artist fixed effects and we cluster standard errors by artist.[29] In these DSP regressions we introduce θ_K to control for country effects.

[27] Following Pollock et al. (2010), we refer to these recordings as being in the public domain, although the composition is still protected by copyright.

[28] Or 2019 for some robustness checks.

[29] Our results are also robust to standard errors clustered by original release year.

Although in the Spotify data we cannot observe the year of expiry as we do in the panel datasets described above, we do estimate specifications with fixed effects for original release year interacted with the UK coefficient. This allows us to observe whether there is a discontinuity in availability for the UK relative to the US that affects starting with songs recorded in 1963. If there are effects of copyright on Spotify availability, we would expect to see greater availability in the UK than in the US for songs recorded in 1959-1962, with a drop in availability 1963 that is sustained for subsequent years.

We exploit variation in copyright status between the US and UK for estimation of digital availability due to the unavailability of a panel measuring digital streaming availability before and after copyright expiry. Relying on cross-country data to form a control group is a strong assumption given the economics of both contexts, so rather than estimating counterfactual outcomes with US observations we prefer to derive variation in copyright status from a natural experiment and staggered lapse into the public domain when able. Nonetheless, robustness results from triple-difference estimators for releases and performances are available in Tables 4 and 5 as described in the subsequent section. These results confirm a positive impact of copyright expiry on UK releases relative to a US control group with no expiry. Similarly, they confirm a negative impact of expiry on UK performances relative to a US control group.

4.2 Results on Re-Releases: Physical Releases

Figure 2 displays the mean number of re-releases by age, for songs released prior to 1963. There appears to be a discontinuous jump in releases after age 50, when these recordings lapse into the public domain. However, there also appears to be a trend change in the trajectory of releases after approximately age 30, perhaps due to fluctuations in demand for physical releases over time. To account for this, we also display in Figure 2 the residualized mean releases, computed from the residuals of a regression of the outcome variable on dummies for year t, which controls for variation over time in the overall number of physical releases. The residualized means have a much flatter trajectory until the large jump in availability upon rights expiry after age 50.

Table 4 displays the baseline results from regressions in which the dependent variable is the number of UK re-releases of song i in year t. Standard errors are clustered by artist in specifications that control for artist fixed effects, and by song in specifications that control for song fixed effects. Because these are Poisson regression coefficients, the percent increase in releases from public domain status is $e^{\alpha_1} - 1$. Column (1) includes controls for year of release and song, and the Poisson coefficient estimate on the PD dummy variable is 1.341 with a standard error of (0.024), which implies a highly statistically significant increase of 282% in the number of re-releases after recording copyright expires. Controlling for the age of the song (number of years since original release), year of release, and artist in column (2) increases the estimate to 1.413 (standard error of 0.250, percentage change of 311%). Column (3) removes artist controls and instead controls for song fixed effects, which increases the estimate to 1.488 (standard error of 0.067, percentage change of 343%).[30]

Column (4) introduces a triple-difference regression in which US releases of song i in year t act as a control for the UK re-releases of the same song. We continue to estimate a large positive coefficient on the public domain dummy, consistent with our expectation and confirming the importance of copyright as the main explanation for our findings. Column (5) is an OLS regression with the dependent variable capturing the share of releases from reissue labels, the coefficient point estimate of 0.208 (standard error of 0.027) shows that the share of new releases by reissue labels increases by approximately 23% post copyright expiry.

Column (6) presents an event study regression for the years before and after copyright expiry, with the year of expiry as the omitted category.[31] This demonstrates a sharp increase in availability upon copyright expiry, also shown in Figure 5. In an additional robustness check not reported in the table, we estimated a regression-discontinuity model. Results are robust to estimating a regression discontinuity specification with a linear trend in years to expiry, a linear trend in years post expiry, and a dummy for post expiry.[32]

The coefficients in Column (6) increase up to the final year of copyright protection (T=0), followed by a large jump at the first year in public domain (T=1). Because copyright terms were extended by 20 years in 2013, songs released in 1963 or later have their "years to expiry" variable reset in 2013, so that songs released in 1966 (for example) are never less than 4 years from expiry, and songs released in 1965 reach a maximum of 3 years prior to expiry. Thus songs in the -1 category have systematically older release dates than songs in the -2 category, and so on. The addition of a linear control for year of original release (Column 7) diminishes the pre-trend observed in Column (6). Column (8) controls for a linear age term and a differential trend in age for songs released before 1963, effectively eliminating the pre-trend. Figure 5 displays the event-study coefficients found in Columns (7) and (8) of Table 4.

Our results are based on a comparison of songs first released before 1963, which enter the public domain at age 50, with songs released later that remain under copyright. If there is a difference in age trends between these groups, our estimates will be biased. To establish robustness of our results to the possibility that trends may not be parallel among the pre-1963 and later songs, we estimate several different specifications in Appendix Table 12. The results are described in detail in the Appendix (see the section titled "Identification and Robustness to Parallel Trend Assumption"). To summarize, the increase in releases at copyright expiry is robust to limiting the sample to songs with original release years before 1963, to narrowing the sample to songs released between 1960 and 1965 (or a still narrower range), and to allowing for pre-1963 songs to have a differential trend in age. Column (8) of Table 4 displays the event study coefficients after controlling for original release

[30]It is not possible to control for age, year, original release year and artist fixed effects because the artist fixed effect is collinear with the original year effects.

[31]Four years pre- and post- expiry are included in the table. The regression includes dummies for all years prior to expiry, and these coefficients are omitted from the table but are available upon request.

[32]This specification, estimated with OLS, yields a coefficient of 0.299 (s.e. 0.011) on the dummy for post copyright expiry, or a 453% increase in releases at expiry.

year and a differential age trend. Following the suggestions of Rambachan and Roth (2020), we also perform a sensitivity analysis which shows that any kink in the differential trend at age 50 would have to be implausibly large to explain our results (see the Appendix).

Appendix Table 8 contains several addditional robustness checks on the main result.[33] To address potential concerns that the results are driven by a handful of extremely successful artists, in columns (1) and (2) we exclude the top 5 artists in the sample in terms of number of releases (The Kinks, Elvis Presley, The Beatles, Frank Sinatra, and Jerry Lee Lewis). Results are once again very similar to the equivalent columns in Table 4. In Columns (3) and (4) we remove the age restriction for our estimation sample and include songs more than 54 years old. Results are similar to the equivalent specifications in Table 4. Columns (5) and (6) of Table 8 present OLS estimates where the dependent variable is the log of the number of re-releases (+1) of song i in year t in the US market. As with the Poisson estimates, we observe a large and significant increase in availability upon copyright expiry in the linear model. In Column (7) of Table 8 the dependent variable is the release count in the US of song i in year t, and shows a small and statistically insignificant effect of public domain status. We restricted the data for the placebo regression to the more homogeneous sample of songs first released between 1960-1965. This result is consistent with our expectation and confirms the importance of copyright as the main explanation for our findings.

Due to the fact that releases and performances of a song may be an infrequent event, we collapse the panel data to the track-five year level in Table 15 of the Appendix. Releases and performances were aggregated into five-year bins, from 1960-1964, 1965-1969, etc. Average song ages were taken within each year bin, and the average age was further discretized into five-year age bins. In this aggregated sample, the mean number of releases in a five-year bin is 0.315, and the mean number of performances is 1.008. We regressed these five-year aggregated counts of the outcome variable on a linear control for five-year age bin, five-year fixed effects, and song or artist fixed effects. Results are robust to collapsing the data in this way, with effect sizes similar to Tables 4 and 5.

4.3 Set List Results

Figure 3 shows average song performances by age, for songs released prior to 1963. Residualized performances in Figure 3 account for the substantial variation over time in the overall number of song performances, and show a decrease in performances after age 50, when pre-1963 sound recordings entered the public domain.

Results on public performances of songs are found in Table 5. Data are restricted to years in which the artist was active.[34] The regression model is similar to the one described in Section 4.1, where the unit of observation is the number of performances of song i by artist j in year t with song age A. The estimation method is a fixed effects Poisson regression, and fixed effects for year,

[33] The main results in Table 4 were estimated using OLS and Logit, with very similar results, always implying a large and significant increase in the number of re-releases after the expiry of recording copyright.

[34] Artist active years are gathered from MusicBrainz.

age, and artist or song are included depending on the specification. Standard errors are clustered by either artist or song.

In columns (1)-(3) of Table 5, the dependent variable is the total number of concert performances in the United Kingdom. Column (1) includes controls for year of release and song, and the coefficient estimate of the public domain effect is -1.401 with a standard error of (0.257), which implies a statistically significant decrease of 75.4% in the number of performances after recording copyright expires.[35] Adding artist and age fixed effects, while removing song fixed effects, in column (2) increases the magnitude of the estimated effect to -1.788, or a 83.3% reduction in the number of performances, significant at the 1% level.[36] Column (3) includes age, year, and song fixed effects and results in a public domain coefficient of -1.898, or an 85.0% reduction in the number of performances, significant at the 1% level.

Column (4) in Table 5 reports the results of a triple-differences regression in which performances in the UK are compared to a control group of performances in the US. There continues to be a negative and significant coefficient on the public domain dummy in this alternative specification, although the coefficient is smaller in magnitude than in Column (3), and implies a 57% reduction in performances. It is possible that this smaller effect reflects spillovers from UK set list choices into US performances.[37] However, it is worth noting that the smaller coefficient estimate from the triple-difference specification is similar to the coefficients in the regressions on UK-only data that account for possible deviations from parallel trends in Table 13, discussed below.

Column (5) of Table 5 displays coefficients from an event study prior to and post copyright expiry, and a clear decline post expiry is apparent. Column (6) estimates the same specification, with the addition of controls for linear year of original release and a differential trend in age for songs released before 1963. Coefficients are similar after the addition of these controls. Event studies are shown graphically in Figure 5.

The number of performances shows an increase prior to copyright expiry driven by The Searchers, a UK group who gave large numbers of performances on their 50th anniversary tour. The Searchers had the highest average number of UK performances in this period, and no songs recorded before 1963. Using the method used to calculate years to/from copyright expiry in Column 5 of Table 5, this group's performances in 2012 are classified as occurring 1 year before expiry because the copyright term extension had not yet taken place. However, in hindsight of the term extension they were 21 years from expiry. When we calculate years to expiry using only songs recorded before

[35] Results are similar when only controlling for age or for year effects individually.

[36] To ensure the results are not driven by the performance decisions of the most popular artists, we also tried excluding the top 5 artists measured in terms of the number of performances (The Beach Boys, Bob Dylan, Frank Sinatra, The Who, and The Rolling Stones). Estimation results are comparable to the equivalent regressions in column (2).

[37] Artists may tour both the UK and the US in the same year but cannot be in both locations at once. In addition, artists who optimize set lists for the UK market may end up playing some of the same songs in the US because they have rehearsed and played them on the same tour. Therefore the decision to perform a particular song in the UK may affect the decision to play it in the US. For this reason, the use of the US performances as a control group for performances may violate the stable unit treatment value assumption (SUTVA), and be less ideal than using the US controls for re-releases (described previously).

1963, the ramp-up is not apparent, but the post-expiry decline remains statistically significant. This specification is included in Column (2) of Table 13.

Table 9 shows additional robustness checks on Setlist results. Columns (1)-(2) use ordinary least squares as the estimation method and the natural logarithm of performances (+1) as the dependent variable. Columns (3)-(4) show results from the main specifications, including data on songs that are more than 54 years old. If the expiry of copyright has an effect on the artist's decision to tour, the results in Table 5 combine this effect with any potential effect on the decision of which songs to perform. Columns (5)-(6) are conditioned on the artist touring in year t, and therefore isolate the choice of songs. Columns (7)-(8) exclude the top 5 artists measured in terms of the number of performances.[38] Column (9) uses the count of US performances as the dependent variable and shows a small and insignificant effect of public domain status. The sample in this specification is limited to original release years between 1960-1965 in order to be conservative and most consistent with our matched sample style event study specifications. In an additional robustness check not reported in the table, we estimated a regression-discontinuity model. Results are robust to estimating a regression discontinuity specification with a linear trend in years to expiry, a linear trend in years post expiry, and a dummy for years post expiry as well as year and artist or song fixed effects.[39]

These results do not appear to reflect simple reallocation across songs within a fixed number of performances. Appendix Table 11 presents results collapsed to the artist-year level which analyze the probability of having any performances in the UK or US in a given year. These results are described in the appendix, and suggest that artists do not merely shift from public domain songs to copyrighted songs: for a given age profile, having a higher percentage of songs protected by copyright in the UK appears to increase the probability that an artist will perform live in the UK. This increase in touring propensity could lead to an increase in live performances even for the inframarginal songs.

It is worth noting that the nature of our data and preferred econometric approach limits our ability to identify the impact of copyright status on performances, given the length of copyright relative to typical music industry careers.[40] This is because we control for age and artist effects in our regression, and given these controls can only identify the impact of copyright status in the years before and after 2012, when a given artist will have songs of approximately the same age both on- and off-copyright. The number of artists who recorded songs both before and after 1963 who are

[38]To estimate the effect separately for the most popular songs in our dataset, we created a dummy equal to 1 if the number of re-releases of the song is above the 99th percentile in our dataset, and then interacted that dummy with the public domain dummy. In a specification that controls for age, year and artist fixed effects, the coefficient on this interaction term is -0.653 (with a standard error of 0.366) and the coefficient on the public domain dummy is -1.72 (with a standard error of 0.584). This finding is similar when controlling for song fixed effects.

[39]This specification, estimated linearly, estimates a coefficient of -0.733 (s.e. 0.237) on the dummy for post copyright expiry, or a 300% reduction in performances at expiry.

[40]There are 26 artists who have observations in the setlist data in which at least one song is in the public domain, and therefore have the variation to capture an artist-specific public domain effect. We explored the possibility of including performances from more "obscure" artists using data collected for Spotify analysis. Of the 41 "obscure" artists, none are: found in the setlist.fm data, have tracks released between 1960-1975, have songs both pre-1963 and post-1963, are active in 2012, and are primarily UK based.

still performing after 2012 is relatively small.[41]

It is clear that some artists prefer to play copyrighted songs over public domain ones (all else equal), both because it is apparent in the data and because Cliff Richard has stated that he pursues this strategy (see section 2.2). However, it is unclear how important this effect is for recording artists in general. This is because only a small number of artists from the pre-1963 cohort were still performing when their songs entered the public domain. This means that we do not directly observe a within-artist decline in performances of public domain songs for very many artists, making it hard to tell how important the "Cliff Richard effect" is as a phenomenon for live performances in general. We can, however, use as a counterfactual the performances of songs aged more than fifty from the later cohort, who recorded songs in 1963 or later and had their copyrights extended. Among this group, we observe resilient performances or even an increase around age 50 (see for example, the Searchers' aforementioned 50th anniversary tour). The comparison of this group and the older cohort, who do not show significant performances after age 50, helps us identify the effect of extended copyright terms on performances.

The set list event study in Table 5 and Figure 5 showed no evidence of a downward pre-trend that could potentially explain the decline in performances post copyright expiry. Nonetheless, we estimate similar specifications to those in Table 12 on the set list sample and describe them in Appendix Table 13. We continue to estimate a robust negative and significant effect after restricting the sample and incorporating differential trends by original release year. However, the estimated effect is smaller in the sample restricted to songs first released before 1963 (in which we identify the effect of copyright expiry solely from the discontinuity in copyright status at age 50) than in the full sample in which we can control for age effects. The magnitude of the effect of public domain status on performances in the restricted sample is -63% (Column 1 of Table 13). In the full sample with age effects, it is -85% (Column 3 Table 5). Other estimates which control for differential trends or narrow the sample years range from -72% (Column 6, Table 13) to -84% (Column 7 Table 13).

The increase in the magnitude of the effect after controlling for age in the performance regressions is consistent with a 50th anniversary effect that biases estimates toward zero. This is plausible, as many of the artists in our sample went on 50th anniversary tours. The positive effect of the 50th anniversary on performances makes it difficult to identify the impact of copyright expiry in the pre-1963 sample, in which all songs enter the public domain at age 50, and will cause those estimates to understate the negative effect of copyright expiry on performances. The increase in the magnitude of the negative effect in the full sample shows the importance of controlling for age effects and exploiting data from around the time of the term extension to obtain identification.

[41] There are 31 artists still touring after 2012. Of these, 13 artists have at least one public-domain recording. We considered expanding the sample to include more artists. However, we did not find sufficient data on the performances of more obscure artists on Setlist.fm in the post-2012 period to warrant collecting a larger sample.

4.4 Availability on Spotify

In this section, we estimate the impact of copyright protection on song availability on Spotify. Figure 6 depicts the lack of a public domain effect on DSP availability with a difference-in-differences model, wherein the availability of songs in the UK market is compared to their availability in the US geographic market. There does not appear to be any statistically significant difference in availability in the UK for songs released prior to 1963, those in the UK public domain, compared to availability in the US.

We find that sound recordings that have entered the public domain are no more likely to be available on Spotify than recordings still protected by copyright. Table 6 compares the availability in the UK geographic market for songs originally released before 1963 with the availability of the same songs in the US geographic market. Our results suggest that sound recordings released before 1963, (and hence in the UK public domain), are approximately 1% more likely to be available for streaming in the UK than in the US (where the sound recording has not fallen into the public domain) but that this difference is statistically insignificant at the 5% level.[42] This small and statistically insignificant difference is of limited economic significance when compared to our results covering copyrights impact on physical releases.

A possible caveat to this finding is that we have restricted our analysis to artists that appeared on top twenty weekly album charts in the early sixties. In order to address the effect that public domain status may have on less prominent (or "obscure") artists, we collected data on artists that have only one song in weekly top singles charts, and whose song appears on the charts for two or fewer weeks.[43] We incorporate these artists into our population and estimate the regression model in column (5) to capture the effect of public domain status on availability for these songs by obscure artists. The interaction of $Obscure * Pre63 * D_k$ captures the additional effect that public domain status has on these obscure artists. We find that sound recordings from obscure artists that have entered the public domain are approximately 4.3% more likely to be available for streaming in the UK than in the US but that this difference is statistically insignificant at the 5% level. This implies a lack of difference in effect for obscure artists compared to non-obscure artists. This result may be due to DSPs employing a blanket license strategy in which the entire catalog of a label/distributor is negotiated for, leading to a lack of differentiation between the major artists and the obscure ones.[44]

A number of robustness checks on the Spotify dataset are presented in Table 10. Column (1) presents a naive regression, controlling for song and country fixed effects. Columns (2)-(3) present specifications with different combinations of fixed effects on the population that includes the obscure songs from the singles charts referenced above. Column (4) excludes any song that is unavailable in both the UK and the US, to confirm that our results are not driven by errors in our song matching

[42] Recordings made before 1972 are not covered by federal copyright law, but rather by state law, which according to Brooks (2005a) implies that these recordings will enter the public domain in 2067.

[43] Approximately 15% of songs within the singles charts appear for two or fewer weeks.

[44] According to CD Baby, negotiations occur in this blanket format, rather than artist by artist.

process.[45] Column (5) drops the top 5 artists, Column (6) estimates the regression on a restricted sample of songs used in the analysis of Amazon data (see below), and Column (7) excludes any songs above the 90th percentile of age (songs recorded before 1957). Results are robust to these various specifications.

It is possible that some of the most obscure songs will not appear on Spotify due to difficulties in locating rightsholders (the "orphan works" problem). Although we did not identify any significant difference in the effects of copyright on availability at different popularity levels in our sample (as seen above), it is possible that this occurs among songs that are even more obscure than the most obscure songs in our sample (e.g. those that did not appear on the charts) and were necessarily omitted from our analysis as a result of our sample construction. To address this possibility, we collected an auxiliary list of performers appearing on the BBC radio program *The Saturday Club* between 1958 and 1969. Analysis of this sample is described in the appendix. To summarize, we find some evidence that suggests that copyright could limit availability for more obscure songs, however some patterns in the data suggest confounding factors besides copyright may explain the results. More research is needed to understand the effects of copyright on digital availability for the most obscure songs and artists. This question remains exceedingly difficult, however, as the most obscure artists from past vintages are unlikely to be recorded in any digital database.

4.4.1 Availability on Amazon

One question raised by the digital streaming platform results is whether they are a result of digital distribution per se, or of the specific DSP business model, with its comprehensive licensing of whole catalogs. To answer this question, we collected data measuring the availability on Amazon Music of a random sub-sample of the songs in our database (stratfied by recording year). Although Amazon has a streaming service, it also sells individual à la carte digital downloads of songs on both the US site (Amazon.com) the UK site (Amazon.co.uk). We collected data on the number of versions available on each site in February 2021.[46] The methods and results are described in detail in the appendix and results are shown in Table 7. In summary, we perform the same analysis performed on the Spotify data, in which we compare availability in the US and the UK across the pre- and post-1963 time periods.[47] We find a dramatic decline in the number of MP3 versions available for sale on Amazon.co.uk for tracks recorded after 1963, and hence still under copyright in the UK, relative to tracks recorded before. We do not observe a similar decline in availability of the same tracks on Amazon.com (see Appendix Figure 9). Estimating a difference-in-differences regression that compares the difference in availability for songs recorded before and after 1963 in the UK relative

[45] There are likely to be some songs that are available on Spotify but which we incorrectly classify as unavailable in both the UK and the US as a result of the difficulty of matching such a large number of song-artist combinations. If there were any bias introduced by false-negative matches, we would expect results in Column (4) of Table 10 to be substantially different from the results in Table 6. The similarity of the coefficients confirms that this bias is unlikely to be a problem.

[46] In a previous draft, we analyzed data on a different random sample not stratified by year, and obtained very similar results.

[47] Additionally, we conduct the Spotify analysis described above, limited to this random sample, and find very similar results.

to the US, we estimate a 252-398% increase in availability associated with public domain status, significant at the 1% level. Graphically, we observe a sharp change in availability at 1963 similar to the jump observed for physical releases (see Figure 6).[48] Controlling for a differential trend in age reduces the coefficient to 1.259 (standard error 0.087), or a 252.2% increase in availability in the UK relative to songs still under copyright. Thus, when songs are sold in an online marketplace that mimics the "record store" model, entry by low-cost reissue labels drives up availability when recording copyrights expire. Though the typical Spotify listener may be less interested in the artists from the 1960s than the typical purchaser of CDs, we have no reason to suspect a dramatic shift in demand for songs from 1962 versus 1964, for example. Additionally, these Amazon results confirm the idea that it is the digital streaming platform business model, rather than digitization of music distribution per se, which eliminates the effect of copyright expiry on availability.

It is worth noting that, while availability is relatively unaffected by public domain status, actual consumption of songs may still be affected if Spotify privileges public domain songs in promotion. Since Spotify does not have to pay royalties for the recording copyright for public domain songs, they may have an incentive to promote consumption of them over copyrighted songs. If Spotify privileges songs with expired recording copyrights on suggested playlists, we may see higher consumption of these songs than would otherwise be observed. However, we currently do not have access to data on usage patterns for these songs.

4.5 Discussion and Welfare

4.5.1 Results in Context

We have shown that prior findings of an increase in books in print upon copyright expiry (Reimers (2018); Heald (2008b)) correspond well to the effects of recording copyright on song re-releases in the historically dominant physical distribution channel. However, results differ in two separate channels which have come to account for the majority of revenues for the music industry. Our findings should be interpreted in the context of the ongoing evolution of the music industry in order to understand the relative importance of each set of results. In other words, should copyright terms be extended more readily in a world in which DSPs have become predominant? Moreover, although we estimate a positive effect of copyright on live performances in our sample, does this finding generalize beyond our setting, and how important is it for the average music consumer?

The answer to the former question depends in part on the resilience of the market for physical releases. CDs, LPs and other physical formats accounted for over 60% of label revenues in the UK in 2010 and 32% of record labels' revenues in 2017 with total revenues of £310.5M. Sales of vinyl albums (LPs) have grown substantially to account for 5.7% of label revenues by 2017. Starting from 2.9% of revenues in 2010, streaming (subscription, ad-supported and video) generated £388.8M in revenues for labels in 2017, or 40.3% of all label revenues, and 58.9% (or £628M) by 2019 (British

[48] This result is also consistent with what Reimers (2018) found for e-books.

Phonographic Industry (2019)). Gross Value Added from live performances in the UK in 2017 was £991M (UK Music (2019)).

Although the market share of CDs and LPs has fallen substantially, these formats still comprised approximately 20% of the UK market in 2019 (British Phonographic Industry (2019)), and sales of LPs were growing substantially.[49] CDs remain particularly popular in certain sub-markets of the industry, for example among purchasers of greatest hits albums, a category highly relevant in this study (Knopper (2018)). In addition, results on a sub-sample of digital downloads from Amazon Music suggest that the increase in availability observed for physical formats is also seen in the availability of digital downloads (due to the download platform's similarity to a "retail" distribution model). For this reason, our results on availability in physical formats remain relevant even as digital distribution rises in prominence. Moreover, distribution channels that follow the "retail" model may be especially relevant as streaming comprises an even larger share of the market, since physical formats, as durable goods that may be exchanged on secondary markets, provide an alternative for consumers that may moderate the market power of streaming providers.

Comparing the magnitudes of the effects across samples, we estimate that entering into the public domain leads to an increase in physical releases ranging from 163% (Table 4, Column 4) to 343% (Table 4, Column 3). With a mean number of song releases per year of 0.066, implying that the average song will see 0.108-0.226 more releases per year when copyright ends. By contrast, entering into the public domain results in a reduction in the number of performances of a song ranging from 57% (Table 5, Column 4) to 85% (Table 5, Column 3). With a mean number of UK performances per song per year of 0.244, this represents a decline of approximately 0.139-0.207 UK performances per song per year.

It is important to keep in mind that only a minority of artists are still actively performing 40-50 years after their songs were first recorded, and the results should be interpreted in light of the fact that as copyright terms lengthen, the positive impact of copyright protection on the total number of performances will decline as the number of active artists declines. In our sample, 48 artists were still actively performing in 2000, 33 in 2016, and the average total number of performances per year of songs that were 49 years old between 2000 and 2013 was 61.8. However, the increase in re-releases after age 50 in the pre-1963 group of songs is sustained over time (see Figure 2), while performances by most artists in their eighth decades will soon come to an end.

The small number of artists still performing 50 years post recording illustrates the difficulty of estimating the impact of copyright on incentives when copyrights last longer than the performing careers of most artists. The fact that Cliff Richard was so active in lobbying for copyright extension – the Act became colloquially known as "Cliff's Law" – may further suggest that the positive effects of the extension on performances were concentrated among a limited number of still-performing artists, while the negative effects of the extension on re-releases apply to a much larger group of

[49] Musician Jack White has stated that "I definitely believe the next decade is going to be streaming plus vinyl. Streaming in the car and kitchen, vinyl in the living room and the den. Those will be the two formats. And I feel really good about that" (Knopper (2018))

artists whose recordings are still in demand.

Another important consideration is that our sample is based on a set of songs that did not have their copyright terms extended, while the set that did see extensions includes hits by some of the most popular recording artists of all time (The Beatles, Bob Dylan, and The Rolling Stones). It seems reasonable to expect that, were these songs to enter the public domain, we might see larger changes in the number of releases upon expiry.

Our results document mixed effects of copyright expiry on music availability. However, the prior literature has identified several significant potential drawbacks of copyright extensions. These include the maintenance of higher prices and lower availability in traditional retail distribution channels; a lack of variety; and negative effects on cumulative innovation. We discuss these below.

4.5.2 Price-related welfare effects

The effect of copyright on price is perhaps the most salient of these effects. According to Pollock (2009), the study closest to ours in terms of data and legal context, the average price difference between UK pop recordings protected by copyright and those in the public domain is 4-14%. Pollock estimates the deadweight loss associated with copyright in book data (where the price effect is estimated at 5-15%) at 0.1-0.2% of total revenues (both copyrighted and public domain works). By contrast, Reimers's (2018) more detailed structural analysis estimates an average consumer surplus per public-domain title of $9,982, compared to $4,145 in profit per title for copyrighted works. Reimers concludes that "[w]ithout changing incentives for creation, the 1998 Copyright Term Extension was welfare decreasing unless the copyright holder or publishers used the added profits for further innovation." (p. 26)

Our analysis suggests that applying prior estimates of the welfare losses of copyright will be incomplete for two reasons. First, because the act extended copyright during artists' lifetimes, the extension increased incentives for artists to perform copyrighted material. This may have led to an overall increase in live performances for artists with many songs near the 50-year cutoff. Indeed, our analysis in Table 11 suggests that having a larger number of songs in the public domain is associated with a higher propensity to go on tour, controlling for year effects and a quadratic in average song age.[50] While it is clear that our results do not speak to the impact of copyright on incentives to create new recordings, they suggest that copyright encourages artists to increase the supply of complements to existing recordings (live performances). This stands in contrast to prior research which has suggested that extending copyright does not encourage the restoration, maintenance and distribution of pre-existing works (Buccafusco and Heald, 2013). As described above, this increase in live performances is limited due to the advanced age of the artists and the

[50]The extension presumably did not increase performances among younger artists. However, the royalties received from recording copyrights create incentives for any artists receiving those royalties to perform and promote copyrighted songs, regardless of age. Although our results exploit an older sample of songs in which it is possible to identify the effect of copyright, we expect the phenomenon whereby live performances promote sales of copyright-protected recordings to apply equally to younger songs.

relatively small number of artists still performing at that age.

More significantly, the shift to digital streaming platforms with their flat-rate pricing models has attenuated the negative effect of copyright protection on consumer availability.[51] In one sense, this would seem to eliminate the dead-weight loss associated with copyright, since almost all popular music is now available for free on ad-supported platforms such as Spotify and YouTube. However, the advertising that supports the free versions of these services presumably reduces utility for users of the free version of Spotify, who account for approximately half of users, and the other half incur the cost of a monthly subscription fee.[52] Furthermore, the value of these streaming platforms largely flows to consumers with broadband internet access.

Would Spotify's monthly UK subscription price be lower if high-value sixties recordings had entered the public domain after 2013? Royalties paid by Spotify to the UK holders of recording rights to these songs would have fallen if recording copyrights had lapsed, implying a reduction in licensing costs. However, Waldfogel (2020) argues that Spotify maximizes revenues rather than profits. This may stem from the labels' large ownership stakes in Spotify, and their desire to avoid double marginalization (p. 597). Revenue maximization implies that prices are determined by the elasticity of demand, and a reduction in licensing costs would therefore not reduce subscription prices. However, copyright could affect elasticity of demand by limiting competition from entry of CDs released by re-issue labels. Had the term extension of 2013 not occurred, early Beatles, Stones and Dylan recordings might now be available on CD from a variety of reissue labels. The increased availability of some of the most popular recordings of the 20th century on CD could potentially have slowed the transition away from CDs as the dominant format for recorded music.

If one DSP were to become the single dominant music platform, it could have sufficient market power to raise prices with deleterious effects on consumer welfare. In the long-run, consumers that prefer public domain music ("oldies") and who have low valuations for newer music and the DSP's bundle may be excluded from the market.[53] However, inter-platform competition, relatively weak network externalities, cheap substitutes (e.g., CDs), and the outside option of piracy currently constrain any dominant platform's ability to raise prices.

On the supply side of the above scenario, a monopolist DSP would have incentives to maintain access to labels and rightsholders, as it internalizes the value of supply-side competition through the price paid by consumers. Spotify and Apple Music currently provide access to preferred aggregators and limit excess variety on the platform, but this policy provides value by diminishing consumer search costs.[54] Farrell and Weiser (2003) provide a framework for analyzing scenarios

[51] Availability effects similar to CDs are observed in digital downloads on Amazon Music, however the revenue share of digital downloads has declined along with CDs as streaming has become more prevalent.

[52] It is difficult to quantify exactly how much advertising affects utility for the average listener. It is likely to be substantially less than the $9.99 monthly subscription fee for the ad-free service. According to Brynjolfsson et al. (2019), the median consumer of digital music in 2017 valued it at $168/year.

[53] That is, those consumers in the past who greatly benefit when a recording copyright expires and generic recordings proliferate. Of course, bundling may also facilitate transactions for consumers that are willing to pay for the bundle, but otherwise have very low valuations for individual songs.

[54] In physical channels, reissue labels also provided value by marketing unique arrangements of public domain music. In the digital-era, this value is largely internalized by Spotify's playlist features and recommendation engine.

in which a platform may not internalize complementary efficiencies (ICE), and instead act in an anti-competitive manner.

4.5.3 Additional welfare effects

Separate from the price effects of copyright, there are additional welfare implications that have been identified by the prior literature. For example Nagaraj (2018) has identified negative effects on cumulative innovation on Wikipedia; Nagaraj and Reimers (2020) find that free digital access increases sales of books, and Biasi and Moser find negative effects of copyright on knowledge diffusion in science. Simcoe and Watson (2019) find that stronger enforcement of copyrights preventing sampling of songs led to less reuse but also more dispersion in the distribution of sampled work. Given the findings of this prior research, we might expect that copyrighted recordings in our sample are used less in derivative works such as music sampling, films or television shows. Copyright term extensions may thus reduce variety or cumulative artistic innovation by raising the costs of licensing music for use in films and television programs (Heald, 2014a). The cost of rights to include Beatles compositions in the recent film *Yesterday* was reported by Billboard at $10 million, or 40% of the total budget of the film.[55] Additionally, rights holders can be reluctant to grant permission for re-use if there is concern that the derivative work will reduce the value of the underlying right.[56]

Finally, welfare considerations must acknowledge the consumer frictions and switching costs that accompany technological change. For example, the technology used to consume recorded music when the albums of the early 1960s were first released is essentially obsolete today. In a random sample of 1,500 American recordings released between 1890 and 1964, Brooks (2005b) finds that 65% of historic recordings are not available to listeners because they are not reissued by rights holders and because "the physical barriers created by recording technologies change often and have rendered most such recordings accessible only through obsolescent technologies usually found only in special institutions" (p. 14).

Moreover, recent research on book sales has suggested that consumers tend to be wedded to their preferred distribution channel, whether physical or digital (Chen et al., 2018, p. 11) and that, if a song is not available in CD format, consumers may choose a different song available on CD rather than switching to Spotify. This suggests that, if old music is not reissued on CD due to extended copyrights, it may fade into obscurity.

[55] The Long and Winding Road to 'Yesterday,' a Film Full of Beatles Music, by Itzkoff, Dave. The New York Times, International edition; New York, 01 July 2019.

[56] "It has to be the right product, it has to be the right film, it has to be the right brand," he said. "We don't want to be detrimental to the group or their legacy in any way. Its about making sure that every part of it is right." (Itzkoff 2019).

5 Conclusion

We obtain mixed results on the effect of recording copyright term on the supply of music: when a song enters the public domain, there are more reissues of that song, but songs with recordings in the public domain are performed less often in concert. This suggests that when artists are living at the time of a copyright term extension, the negative supply effects of the extension on re-releases may be partially offset by a temporary positive supply response in live performance. The practical importance of this effect depends on the number of artists still living and performing in concert at the time of the extension, and the remaining lifespan of these artists. Given heterogeneity in copyright terms across countries, our results may inform debates surrounding the appropriate duration of rights, the transition of content industries to bundled subscription platforms (e.g., Kindle Unlimited, Audible), and industries in which rightsholders provide complementary effort to copyright protected works.

Our results regarding the supply of re-releases are consistent with prior studies of copyright's impact on the availability of books (Heald, 2008a; Reimers, 2018). This may not be surprising, given that CDs and books share similar distribution models, in which multiple publishers/labels compete to offer desirable editions/releases, and the expiry of copyright lowers entry barriers for those wishing to offer a low-priced edition/release. The resulting entry lowers prices and increases availability for both books and music distributed on CD.

However, our analysis of digital distribution represents a departure from the prior research on copyright and book title availability. Consumers' desire for near universal access and the high fixed costs of negotiating licenses with record labels have led to the concentration of digital distribution among a small number of large platforms, and the entry by "generic" producers observed in the CD market does not occur. Despite this, we observe no difference in availability on Spotify between public-domain recordings and those remaining under copyright in our main analysis sample, presumably due to the blanket licensing of songs by labels to DSPs.

In the long run, the market's shift away from CDs and towards streaming platforms like Spotify may thus work to moderate the negative welfare implications of copyright term extensions. It is possible that this null effect of copyright on availability is unique to Spotify, but the three major music labels and large aggregators like CDBaby also license their catalogs to smaller competing platforms such as Apple Music, Google Play Music, and Amazon Music's streaming service. There may be significantly greater availability of public domain compositions, through excess variety, on open access platforms such as YouTube operating under a notice-and-takedown regime. Further empirical work may also examine whether digital streaming platforms are biasing consumer recommendations to lower-cost public domain content.[57] This result may be limited to our sample of artists that were relatively well-known during their prime. Analysis of an auxiliary sample is suggestive that copyright may indeed reduce availability on DSPs for more obscure artists, possi-

[57] Bourreau and Gaudin (2019) develop a theoretical model in which a platform provider maximizes profits by biasing recommendations towards low-royalty content.

bly due to the "orphan works" problem. Additional research is needed to definitively determine whether results are similar for more obscure or older recordings such as those considered by Brooks (2005b). However, the available evidence as examined in this paper indicates that digital streaming platforms and live performances have partially offset the negative effects of longer copyright terms on the availability of the most popular music from this period.

References

Aguiar, L., J. Claussen, and C. Peukert (2018). Catch me if you can: Effectiveness and consequences of online copyright enforcement. *Information Systems Research 29*(3), 656–678.

Aguiar, L. and J. Waldfogel (2018). As streaming reaches flood stage, does it stimulate or depress music sales? *International Journal of Industrial Organization 57*, 278–307.

Biasi, B. and P. Moser (2018). Effects of copyrights on science-evidence from the US book republication program. *NBER Working Paper* (24255).

Boudreau, K. J. and A. Hagiu (2009). Platform rules: Multi-sided platforms as regulators. *Platforms, markets and innovation 1*, 163–191.

Bourreau, M. and G. Gaudin (2019). Streaming platform and strategic recommendation bias. *Available at SSRN 3290617*.

British Phonographic Industry (2019). Uk record labels' trade income reaches 1.1 billion in 2019. https://www.bpi.co.uk/news-analysis/uk-record-labels-trade-income-reaches-11-billion-in-2019/.

Brooks, T. (2005a). How copyright law affects reissues of historic recordings: A new study. *ARSC Journal 36*(2), 183–203.

Brooks, T. (2005b). *Survey of reissues of US recordings.* Washington, D.C.: Council on Library and Information Resources and Library of Congress.

Brynjolfsson, E., A. Collis, and F. Eggers (2019, 03). Using massive online choice experiments to measure changes in well-being. *Proceedings of the National Academy of Sciences 116*, 201815663.

Buccafusco, C. and P. J. Heald (2013). Do bad things happen when works enter the public domain?: Empirical tests of copyright term extension. *Berkeley Technology Law Journal 28*(1), 1–43.

Chen, H., Y. J. Hu, and M. D. Smith (2018). The impact of e-book distribution on print sales: Analysis of a natural experiment. *Management Science* (Forthcoming).

Cho, D., M. D. Smith, and R. Telang (2017). An empirical analysis of the frequency and location of concerts in the digital age. *Information Economics and Policy 40*, 41 – 47.

Cohen, J. E., L. P. Loren, R. L. Okediji, and M. A. O'Rourke (2015). *Copyright in a global information economy.* Wolters Kluwer Law & Business.

Correia, S., P. Guimarães, and T. Zylkin (2019). PPMLHDFE: Fast poisson estimation with high-dimensional fixed effects. *arXiv preprint arXiv:1903.01690*.

Danaher, B., Y. Huang, M. D. Smith, and R. Telang (2014). An empirical analysis of digital music bundling strategies. *Management Science 60*(6), 1413–1433.

Danaher, B., M. D. Smith, and R. Telang (2017). Copyright enforcement in the digital age: Empirical evidence and policy implications. *Communications of the ACM 60*(2), 68–75.

DiCola, P. (2013). Money from music: Survey evidence on musicians' revenue and lessons about copyright incentives. *Arizona Law Review 55*, 301.

Dobkin, C., A. Finkelstein, R. Kluender, and M. J. Notowidigdo (2018). The economic consequences of hospital admissions. *American Economic Review 108*(2), 308–52.

Edelman, B., M. Luca, and D. Svirsky (2017). Racial discrimination in the sharing economy: Evidence from a field experiment. *American Economic Journal: Applied Economics 9*, 1–22.

Farrell, J. and P. J. Weiser (2003). Modularity, vertical integration, and open access policies: Towards a convergence of antitrust and regulation in the internet age. *Harvard Journal of Law & Technology 17*, 85–134.

Giorcelli, M. and P. Moser (2016). Copyrights and creativity: Evidence from italian operas. *Available at SSRN 2505776*.

Greenstein, S., J. Lerner, and S. Stern (2013). Digitization, innovation, and copyright: What is the agenda? *Strategic Organization 11*(1), 110–121.

Heald, P. J. (2008a). Property rights and the efficient exploitation of copyrighted works: An empirical analysis of public domain and copyrighted fiction bestsellers. *Minnesota Law Review 92*, 1031–1063.

Heald, P. J. (2008b). Testing the over-and under-exploitation hypothesis: Bestselling musical compositions (1913-32) and their use in cinema (1968-2007). *University of Chicago, Public Law Working Paper* (234).

Heald, P. J. (2014a). How copyright keeps works disappeared. *Journal of Empirical Legal Studies 11*(4), 829–866.

Heald, P. J. (2014b). How notice-and-takedown regimes create markets for music on youtube: An empirical study. *UMKC Law Review 83*(2), 313–328.

Hendricks, K. and A. Sorensen (2009). Information and the skewness of music sales. *Journal of political Economy 117*(2), 324–369.

Knopper, S. (2018). The end of owning music: How cds and downloads died. https://www.rollingstone.com/pro/news/the-end-of-owning-music-how-cds-and-downloads-died-628660/. Accessed: 6/27/2021.

Kretschmer, M. (2011). Private copying and fair compensation: An empirical study of copyright levies in Europe. *UK Intellectual Property Office Research Paper* (2011/9).

Li, X., M. MacGarvie, and P. Moser (2018). Dead poets' property–how does copyright influence price? *The RAND Journal of Economics 49*(1), 181–205.

Luca, M. (2017). Designing online marketplaces: Trust and reputation mechanisms. *Innovation Policy and the Economy 17*, 77–93.

MacGarvie, M. and P. Moser (2015). Copyright and the profitability of authorship: evidence from payments to writers in the romantic period. In S. M. Greenstein, A. Goldfarb, and C. Tucker (Eds.), *Economic Analysis of the Digital Economy*, pp. 357–379. University of Chicago Press.

Marr, B. (2017). The amazing ways Spotify uses big data, ai and machine learning to drive business success. *Forbes*.

Miller, P. (2006). Is this the end of popstar royalty? *The Herald*.

Mortimer, J. H., C. Nosko, and A. Sorensen (2012). Supply responses to digital distribution: Recorded music and live performances. *Information Economics and Policy 24*(1), 3–14.

Nagaraj, A. (2018). Does copyright affect reuse? evidence from google books and wikipedia. *Management Science 64*(7), 3091–3107.

Nagaraj, A. and I. Reimers (2020). Digitization and the demand for physical works: Evidence from the google books project. *SSRN working paper*.

Nicolaou, A. (2018, March). Revenue streams: Spotifys bid to generate a profit. *The Financial Times, New York*.

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

Papies, D. and H. J. van Heerde (2017). The dynamic interplay between recorded music and live concerts: The role of piracy, unbundling, and artist characteristics. *Journal of Marketing 81*(4), 67–87.

Pollock, R. (2009). Forever minus a day? calculating optimal copyright term. *Review of Economic Research on Copyright Issues 6*(1), 35–60.

Pollock, R., P. Stepan, and M. Välimäki (2010). The value of the EU public domain. https://doi.org/10.17863/CAM.5539.

Rambachan, A. and J. Roth (2020). An honest approach to parallel trends. *Working paper*.

Reimers, I. (2018). Copyright and generic entry in book publishing. *American Economic Journal: Microeconomics* (Forthcoming).

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. *The Journal of Law and Economics 49*(1), 29–62.

Rochet, J.-C. and J. Tirole (2003). Platform competition in two-sided markets. *Journal of the european economic association 1*(4), 990–1029.

Shiller, B. and J. Waldfogel (2011). Music for a song: an empirical look at uniform pricing and its alternatives. *The Journal of Industrial Economics 59*(4), 630–660.

Simcoe, T. and J. Watson (2019). Digital sampling, copyright assertion and creative reuse. *Working paper*.

Smith, M. and R. Telang (2016). *Streaming, Sharing, Stealing.* MIT Press.

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

Stanley, B. (2011). Copyright extension: good for cliff and the beatles, bad for the little guys? https://www.theguardian.com/music/2011/sep/15/copyright-extension-cliffs-law-beatles. Accessed: 3/27/2020.

Steele, A. (2018, February). Apple Music on track to overtake Spotify in U.S. subscribers. *The Wall Street Journal*.

Synovitz, R. (2003). EU: Copyright loophole exposes 'golden oldies' of pop music. https://www.rferl.org/a/1101848.html.

Theofilos, S. (2013). A copyright extension for EU recordings. http://www.thembj.org/2013/12/copyright-extension-for-european-sound-recordings/. Accessed 5/23/2017.

UK Intellectual Property Office (2018). Copyright term extension for sound recordings.

UK Music (2019). Music by numbers 2019. https://www.ukmusic.org/research-reports/music-by-numbers-2019/.

Varian, H. R. (2006). Copyright term extension and orphan works. *Industrial and Corporate change 15*(6), 965–980.

Waldfogel, J. (2012). Copyright protection, technological change, and the quality of new products: Evidence from recorded music since Napster. *The Journal of Law and Economics 55*(4), 715–740.

Waldfogel, J. (2020). The welfare effects of spotifys cross-country price discrimination. *Rev Ind Organ 56*, 593–613.

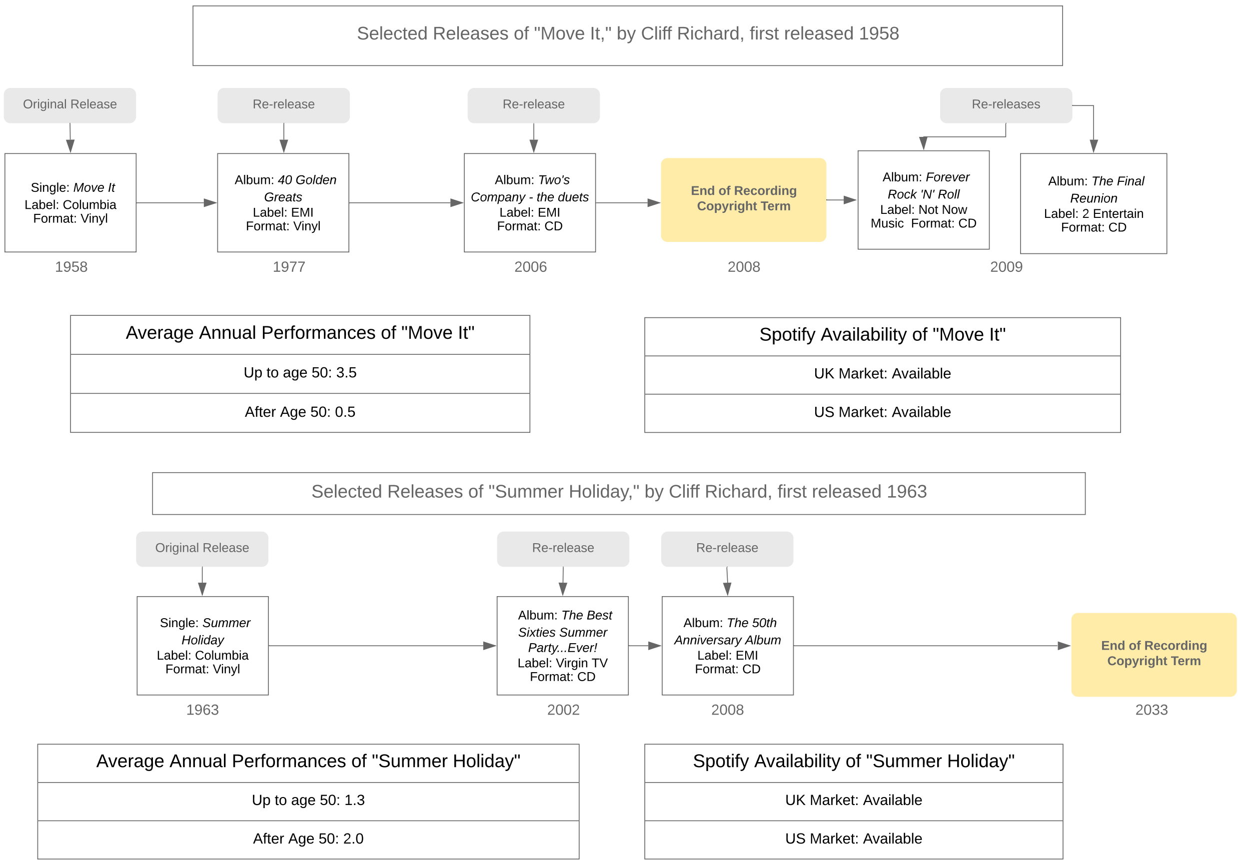
Waterman, D. (2009). *Hollywood's road to riches.* Harvard University Press.

Watson, J. (2018). Copyright and the production of hip-hop music. *Working Paper*.

WIPO (2015). *Guide on Surveying the Economic Contribution of the Copyright-Based Industries.* World Intellectual Property Organization.

Tables and Figures

Figure 1: Copyright Term and Availability



Notes: This figure illustrates the effect of the copyright term extension for two separate songs by Cliff Richard. The top timeline shows selected releases before and after UK copyright expiry for the song "Move It," first released in 1958. The song lapsed into the public domain in 2009, and was subsequently re-released by small independent labels, here depicted in compilation albums by Not Now Music and 2 Entertain. The bottom timeline shows selected releases for the song "Summer Holiday," first released in 1963. This song was re-released on compilation albums by Virgin and EMI in 2002 and 2008. Note that the Columbia and Virgin labels were owned by EMI. Due to the copyright term extension, this song remains under UK copyright protection until the end of 2033. Below the timelines, the average annual number of performances and Spotify availability for each song are displayed.

Figure 2: Raw Effect of Copyright on Physical Releases in the UK

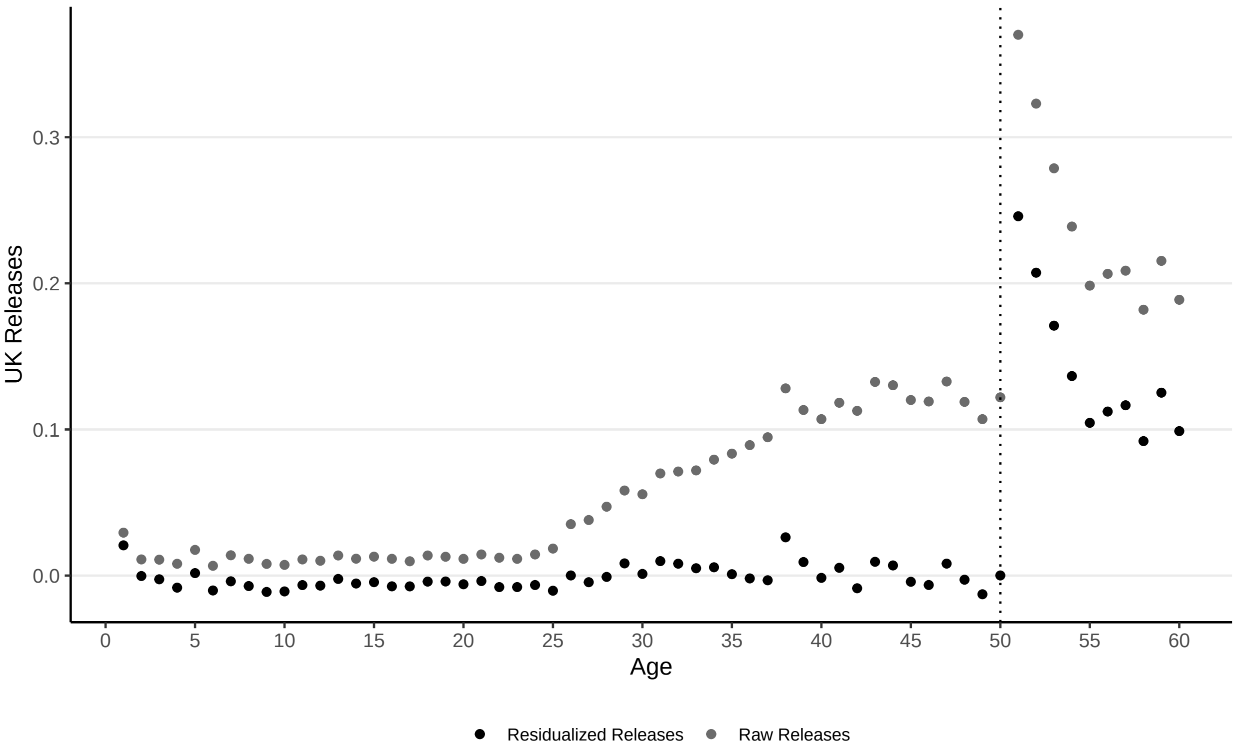
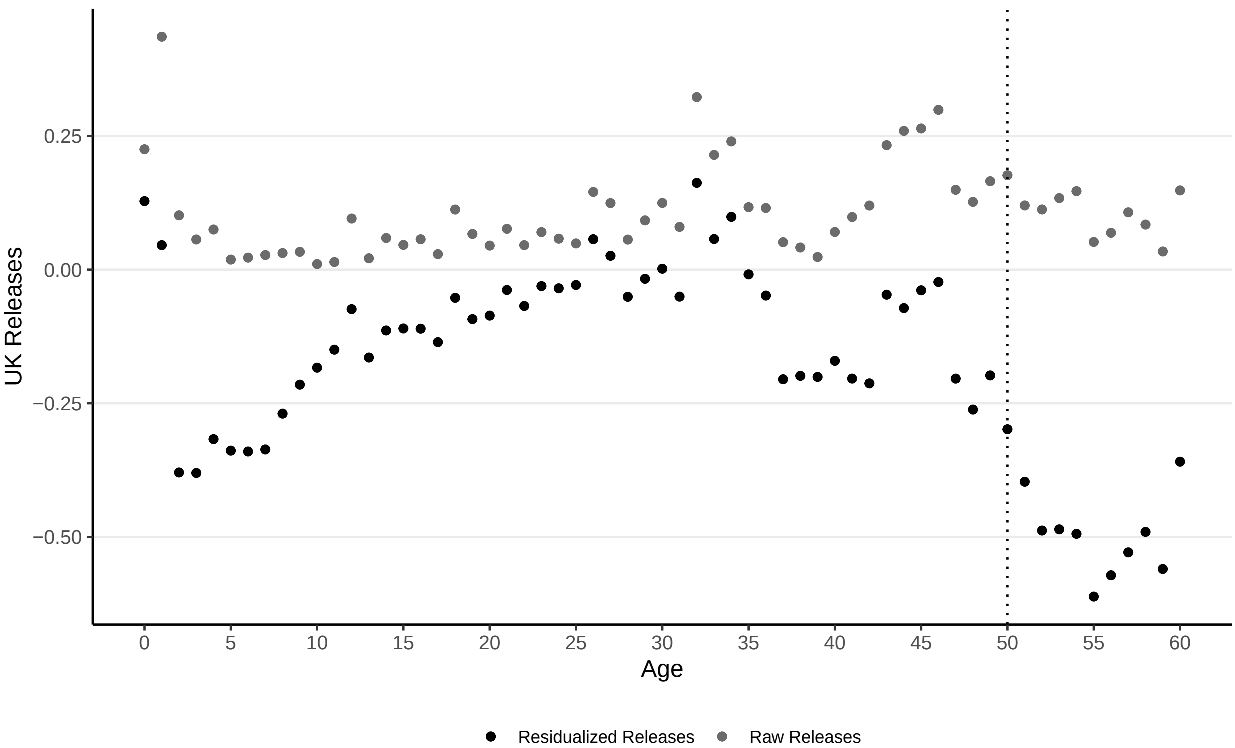
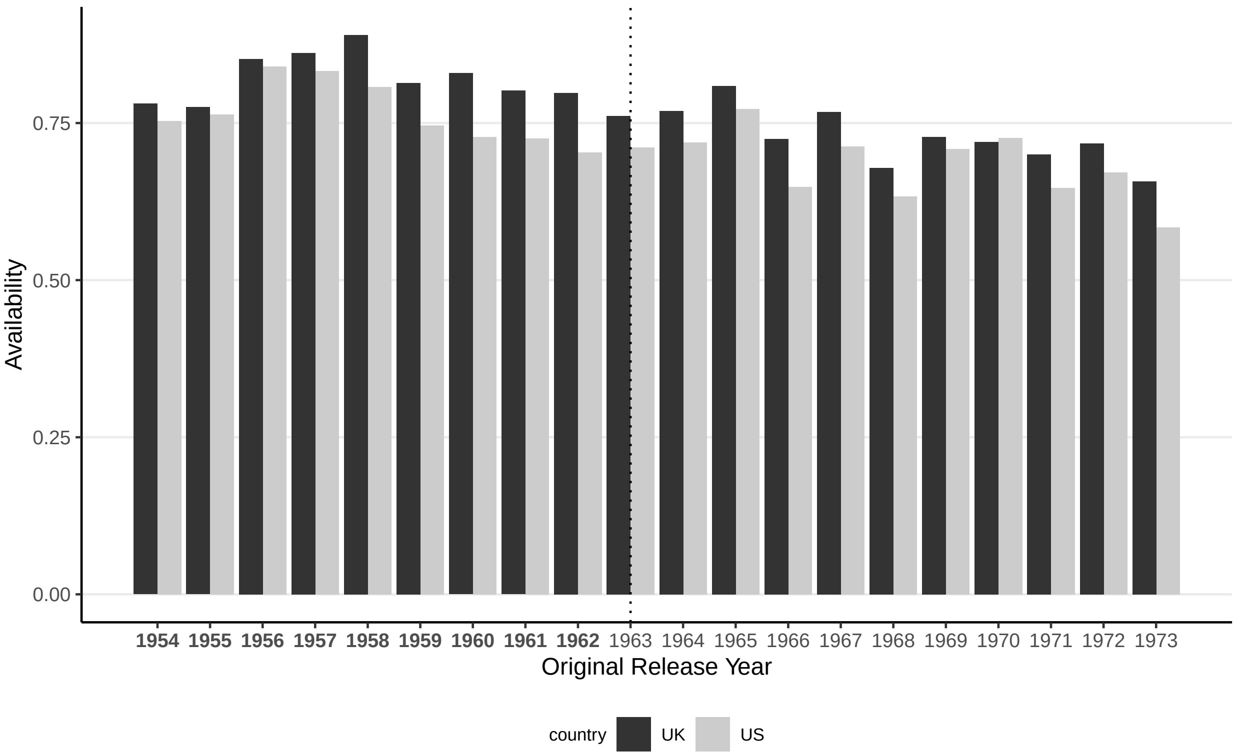


Figure 3: Raw Effect of Copyright on Concert Performances in the UK



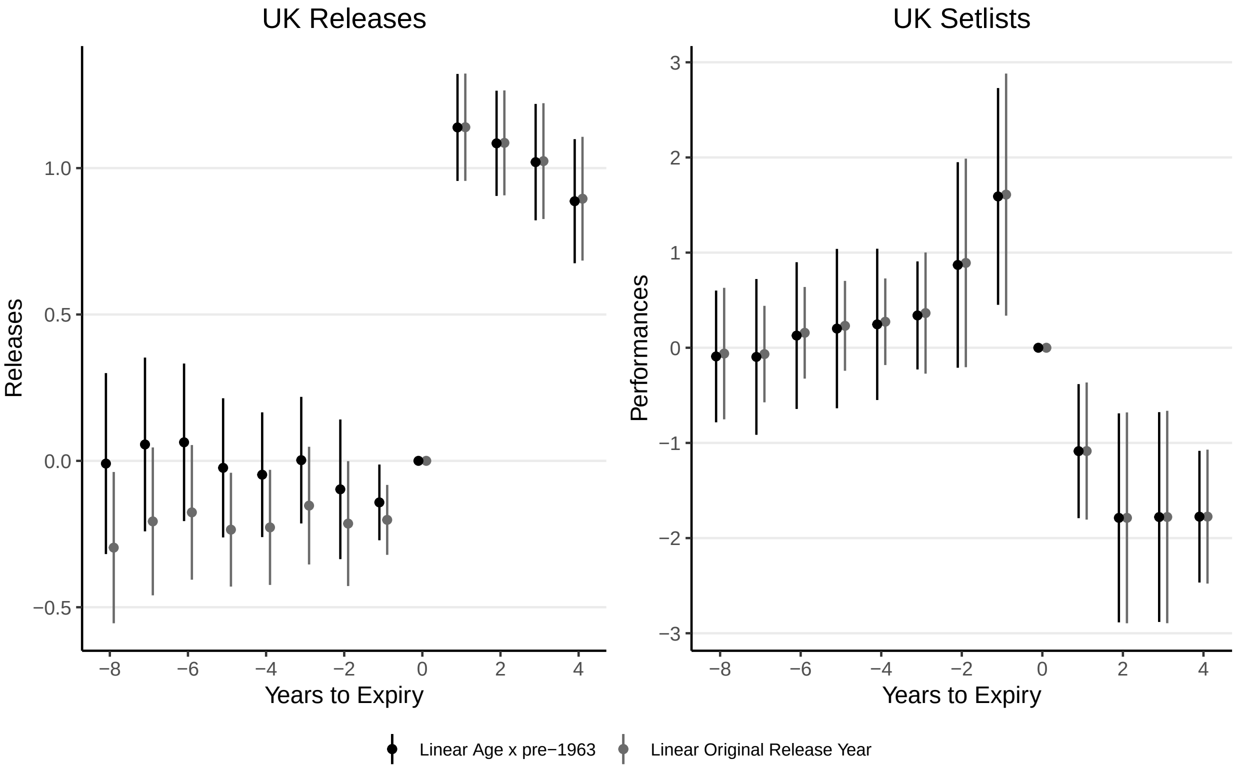
Notes: The top figure shows the average number of physical releases in the UK, by song age. The bottom figure shows the average number of times songs were performed in the UK, by song age. Residualized mean releases are computed from the residuals of a regression of the outcome variable on dummies for year t. The sample is restricted to songs released before 1963, for which the recording copyright term expired at 50 years of age. Songs to the right of the dashed line are in the public domain.

Figure 4: Raw Effect of Copyright on Digital Platform Availability: UK vs US



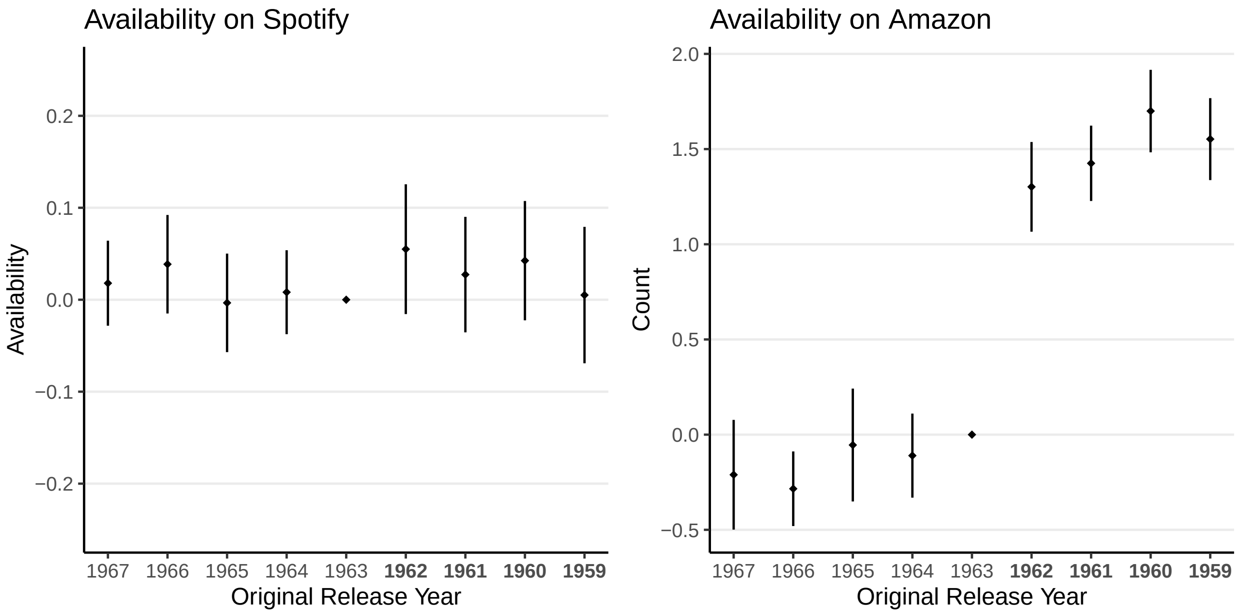
Notes: This figure shows the average availability of songs, by original release year, on the UK and US Spotify platforms. Recordings of songs originally released before 1963 (to the left of the dashed line) are in the public domain.

Figure 5: Time-varying Estimates of the Effect of Copyright on Releases and Live Performances



Notes: The figures display 95% confidence intervals for α_T in the pseudo-poisson regression: $y_{itj} = \alpha_0 + \sum_{T=-50}^{4} \alpha_T + \beta_t + \delta_j + \gamma\, OrigYear_i + \zeta Age \times \mathbb{1}\{OrigYear < 1963\}_{it} + \varepsilon_{itj}$, where $\alpha_{T=-1} = 1$ if song i is 1 year from copyright expiry in year t, and $\alpha_{T=4} = 1$ if song i is 4 years past copyright expiry in year t. The model includes artist fixed effects (δ_j), year fixed effects (β_t), and up to two additional controls. Estimates in gray control for year of original release (γ), while those in black control for year of release as well as a differential trend in age for songs released before 1963 (ζ). The year of copyright expiration ($T = 0$) is the excluded category. The estimation sample includes all observations in which age is less than 55 years, and a full set of dummies for years to/from expiry is included. For illustration purposes the figure is limited to 8 years in the pre-expiry period.

Figure 6: Time-varying Estimates of the Effect of Copyright on Digital Streaming and Retail Availability



Notes: 95% confidence intervals for α_T in the linear regression: $availability_{itj} = \alpha_0 + \sum_T \alpha_T + \gamma_A + \delta_j + \theta_k + \varepsilon_{iAjk}$, where $\alpha_{T=-4} = 1$ for observations of song i in the UK recorded in 1967 (4 years after the cutoff for public domain status, 1963), and $\alpha_{T=4} = 1$ for observations in the UK if song i was recorded in 1959 (4 years from the cutoff for public domain status). The model includes artist fixed effects (δ_j), age fixed effects (γ_A) and country effects (θ_k). US observations are pooled with the excluded category, the last year of copyright protection ($T = 0$ or 1963). For illustration purposes, the figure is limited to four years before and after 1963, but the estimation sample includes all observations in the dataset. For Spotify, the data covers observations with original release years of 1975 and prior, while the Amazon data covers only 1959 to 1967.

Table 1: Summary Statistics on Releases

	(1)			
	Mean	SD	Min	Max
Original Release Year	1963.496	6.624	1923	1975
Year of Release	1990.142	15.091	1960	2017
Age of Track	26.647	14.788	1	54
UK Releases	0.066	0.333	0	12
US Releases	0.078	0.341	0	11
Public Domain	0.030	0.172	0	1
UK Major Label Releases	0.036	0.217	0	8
UK Non-major Label Releases	0.030	0.212	0	9
Observations	922,182			

Notes: Release data from MusicBrainz includes 18,516 songs across 138 artists. "Original Release Year" is the year the song was first released, and "Year of Release" is the year of subsequent re-release. Data is restricted to observations where original release year is 1975 or earlier and age is less than 55. Each song has an observation for each year subsequent to its original release year from 1960 until 2017. Original release year is based on MusicBrainz and is cross-checked with data from Discogs. "Public Domain" is a dummy equal to 1 in years after the end of a song's recording copyright term. We define a release as a "major label" release if the "label type" field on MusicBrainz classifies the release as production, original production, imprint, or holding.

Table 2: Summary Statistics on Set Lists

	(1)			
	Mean	SD	Min	Max
Original Year	1964.199	6.117	1928	1975
Performance Year	1987.145	14.578	1960	2016
Year Performance Count	1.576	7.987	0	178
UK Performance Count	0.244	2.484	0	127
US Performance Count	0.977	5.826	0	122
Touring (= 1 if on tour)	0.633	0.482	0	1
Public Domain	0.012	0.108	0	1
Age	22.945	14.209	0	54
Observations	109,583			

Notes: Setlist data from Setlist.fm includes 2,825 songs across 84 artists. "Original Release Year" is based on MusicBrainz and is cross-checked with data from Discogs. Data is restricted to observations where original release year is 1975 or earlier and for years in which the artist is still active, according to MusicBrainz. Each song has an observation for each year from 1960 until the artist is no longer active (or 2016). "Yearly Performance Count" shows the number of performances of a song in each year, and includes years in which the song was not performed. "Public Domain" is a dummy equal to 1 in years after the end of a song's recording copyright term.

Table 3: Summary Statistics on Online Streaming

	(1) Album Charts				(2) Singles Data			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Original Release Year	1963.449	7.257	1923	1975	1965.375	5.493	1952	1975
Available	0.746	0.435	0	1	0.708	0.455	0	1
UK	0.500	0.500	0	1	0.500	0.500	0	1
pre-1963	0.380	0.485	0	1	0.375	0.484	0	1
Obscure	0.000	0.000	0	0	1.000	0.000	1	1
Observations	37,032				2,686			

Notes: Online streaming data was gathered from Spotify in September 2017, matched to MusicBrainz release data. The Album Charts data in Panel 1 include the same songs and artists found in Table 1. Each song has two observations in our data, one denoting availability in the US market and one denoting availability in the UK market as of 9/2017. Panel 1 contains 18,516 songs by 138 artists in the UK and US markets, for a total of 37,032 observations. Panel 2 includes the additional population of obscure artists used in Column (5) of Table 6, with 1,343 songs by 41 artists and is a cross section as of 7/2019. Data is restricted to observations where original release year is 1975 or earlier.

Table 4: Baseline Results on Releases

	Poisson				*OLS*	*Poisson*		
	(1) UK Releases	(2) UK Releases	(3) UK Releases	(4) UK vs US	(5) ShareNew	(6) UK Releases	(7) UK Releases	(8) UK Releases
Public Domain	1.341***	1.413***	1.488***	0.968***	0.208***			
	(0.024)	(0.250)	(0.067)	(0.026)	(0.027)			
UK				-0.267***				
				(0.013)				
T = -4						-0.500***	-0.227**	-0.047
						(0.103)	(0.100)	(0.109)
T = -3						-0.360***	-0.153	0.002
						(0.106)	(0.103)	(0.110)
T = -2						-0.352***	-0.214**	-0.097
						(0.113)	(0.109)	(0.122)
T = -1						-0.269***	-0.202***	-0.142**
						(0.063)	(0.061)	(0.066)
T = 0						0.000	0.000	0.000
						(.)	(.)	(.)
T = 1						1.129***	1.140***	1.139***
						(0.096)	(0.094)	(0.093)
T = 2						1.093***	1.086***	1.085***
						(0.101)	(0.091)	(0.092)
T = 3						1.045***	1.024***	1.020***
						(0.122)	(0.101)	(0.101)
T = 4						0.879***	0.895***	0.887***
						(0.162)	(0.108)	(0.108)
Age FE	No	Yes	Yes	Yes	Yes	No	No	No
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Artist FE	No	Yes	No	No	No	Yes	Yes	Yes
Song FE	Yes	No	Yes	Yes	Yes	No	No	No
Orig. Yr. Trend	No	No	No	No	No	No	Yes	Yes
Pre-'63 Age Trend	No	No	No	No	No	No	No	Yes
N	572,665	920,439	572,665	1,555,256	42,890	920,439	920,439	920,439

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table displays estimated coefficients from regressions in which the dependent variable is the number of UK releases of song i in year t. Columns (1)-(4) and (6)-(8) are Poisson regressions, column (5) is Ordinary Least Squares (OLS). Age fixed effects control for the number of years since original release. Year fixed effects control for the year of release. Song fixed effects control for the artist-song combination. All specifications control for year FE. Column (1) also controls for song FE. Column (2) controls for age and artist FE, and column (3) controls for age and song FE. Column (4) is a difference-in-differences regression using a Poisson specification comparing the number of releases in the UK to the number of releases in the US before and after songs lapse into the public domain. Column (5) is an OLS regression in which the dependent variable is the share of releases from new labels. Columns (6)-(8) control for time fixed effects before and after copyright expiry, with the year of expiry as the omitted category. Column (7) adds a linear control for original release year. Column (8) controls for year of original release and a differential trend in age for songs released before 1963. Standard errors are clustered at the level of the respective group fixed effect, artist or song.

Table 5: Set List Results

	Poisson					
	(1)	(2)	(3)	(4)	(5)	(6)
	UK Performances	UK Performances	UK Performances	UK vs US	UK Performances	UK Performances
Public Domain	-1.401***	-1.788***	-1.898***	-0.838***		
	(0.257)	(0.608)	(0.364)	(0.246)		
UK				-1.381***		
				(0.079)		
T = -4					0.375	0.336
					(0.249)	(0.478)
T = -3					0.441	0.408
					(0.317)	(0.323)
T = -2					0.944*	0.915*
					(0.532)	(0.536)
T = -1					1.637**	1.612***
					(0.637)	(0.570)
T = 0					0.000	0.000
					(.)	(.)
T = 1					-1.201*	-1.201*
					(0.727)	(0.721)
T = 2					-1.628***	-1.626***
					(0.329)	(0.321)
T = 3					-1.531**	-1.528**
					(0.740)	(0.768)
T = 4					-2.786***	-2.780***
					(1.023)	(1.047)
Age FE	No	Yes	Yes	Yes	No	No
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Artist FE	No	Yes	No	No	Yes	Yes
Song FE	Yes	No	Yes	Yes	No	No
Orig. Yr. Trend	No	No	No	No	No	Yes
Pre-'63 Age Trend	No	No	No	No	No	Yes
N	56,266	97,804	56,266	193,568	97,456	97,456

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table displays estimated coefficients from a Poisson regression where the dependent variable is the number of times a song i was performed in year t in the United Kingdom (columns 1-3 and 5-6). Age fixed effects control for the number of years since original release. Year fixed effects control for the year of release. Song fixed effects control for the artist-song combination. All columns control for year FE. Column (1) also controls for song FE. Column (2) controls for age and artist FE, and column (3) controls for age and song FE. Column (4) shows a DiD regression comparing the number of performances for song i in the UK and US before and after song i lapses into the public domain. Columns (5)-(6) control for time fixed effects before and after copyright expiry, with the year of expiry as the omitted category. Column (5) includes all original release years. Column (6) also controls for linear year of original release and a differential trend in age for songs released before 1963. Standard errors are clustered at the level of the respective group fixed effect, artist or song.

Table 6: Availability on Spotify

	L.P.M				
	(1) Availability	(2) Availability	(3) Availability	(4) Availability	(5) Availability
Public Domain (pre1963 x UK)	0.014	0.014	0.014		0.014
	(0.016)	(0.016)	(0.016)		(0.016)
pre1963	0.068^{**}				
	(0.030)				
T=-4 (1967 x UK)				0.019	
				(0.023)	
T=-3 (1966 x UK)				0.040	
				(0.027)	
T=-2 (1965 x UK)				-0.002	
				(0.027)	
T=-1 (1964 x UK)				0.010	
				(0.023)	
T=0 (1963 x UK)				0.000	
				(.)	
T=1 (1962 x UK)				0.056	
				(0.036)	
T=2 (1961 x UK)				0.030	
				(0.032)	
T=3 (1960 x UK)				0.041	
				(0.033)	
T=4 (1959 x UK)				0.007	
				(0.038)	
UK x Obscure					0.061^{*}
					(0.031)
pre63 x Obscure					0.051
					(0.049)
pre63 x ObscureUK					0.043
					(0.037)
Age FE	No	Yes	Yes	Yes	Yes
Artist FE	No	Yes	No	Yes	Yes
Song FE	No	No	Yes	No	No
Country FE	Yes	Yes	Yes	Yes	Yes
N	37,032	37,032	37,032	19,596	39,718

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table displays coefficients from a linear probability model. In all specifications, the dependent variable is a binary variable that equals 1 if song i is available in geographic market k on the digital music streaming platform Spotify as of September 2017. The estimation sample is restricted to songs with an original release year before 1975. Pre-1963 equals 1 for songs with an original release year prior to 1963. Age fixed effects control for the number of years since original release. Song fixed effects control for the artist-song combination. Country fixed effects control for the US market versus the UK market. All columns control for country FE. Column (2) also controls for age and artist FE, and column (3) controls for age and song FE. Column (4) controls for time fixed effects for number of years before and after the cut-off for public domain status (1963), with 1963 as the omitted category. Column (5) includes data on more obscure artists, who have one song appear in singles charts and whose song is on the charts for two or fewer weeks. Standard errors are clustered by artist.

Table 7: Availability on Amazon

	Poisson					
	(1)	(2)	(3)	(4)	(5)	(6)
	Version Count	Version Count	Version Count	Version Count	Version Count	Drop 1966
Public Domain	1.606***	1.606***	1.606***	1.259***		1.581***
	(0.085)	(0.085)	(0.049)	(0.087)		(0.087)
pre-1963	0.861***					
	(0.113)					
Age × Treated				0.084***		
				(0.018)		
T = -4 (1967 × UK)					-0.211	
					(0.147)	
T = -3 (1966 × UK)					-0.284***	
					(0.100)	
T = -2 (1965 × UK)					-0.055	
					(0.151)	
T = -1 (1964 × UK)					-0.110	
					(0.113)	
T = 0 (1963 × UK)					0.000	
					(.)	
T = 1 (1962 × UK)					1.302***	
					(0.120)	
T = 2 (1961 × UK)					1.425***	
					(0.101)	
T = 3 (1960 × UK)					1.700***	
					(0.110)	
T = 4 (1959 × UK)					1.552***	
					(0.110)	
Age FE	No	Yes	Yes	Yes	Yes	Yes
Artist FE	No	Yes	No	No	Yes	Yes
Song FE	No	No	Yes	Yes	No	No
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
N	3,676	3,644	3,620	3,620	3,644	3,252

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table displays coefficients from a Poisson model. The dependent variable is a count of the number of versions of song i available in geographic market k (UK or US) on Amazon Music as of March 2021. The estimation sample is a random sub-sample of the songs in Table 6, stratified by year of original release. Pre-1963 equals 1 for songs originally released prior to 1963. Age fixed effects control for the number of years since original release. Song fixed effects control for the artist-song combination. Country fixed effects control for the US market versus the UK market. All columns control for country FE. Column (2) also controls for age and artist FE, and column (3) controls for age and song FE. Column (4) controls for age, song and country effects as well as for a linear age term interacted with the UK dummy. Column (5) controls for time fixed effects for number of years after and before the cut-off for public domain status (1963), with 1963 as the omitted category. Column (6) drops songs released in 1966 from the sample as a robustness test due to the statistically significant 1966 coefficient in the Figure 6 event-study. Standard errors are clustered at the level of the respective group fixed effect, artist or song.

Online Appendix

Identification and Robustness to Parallel Trend Assumption

In this section, we describe the robustness of our results to potential violations of the parallel trend assumption. Our preferred specification with age, year and song fixed effects relies on having an "untreated" group of songs that do not enter the public domain at age 50, because this allows us to identify the effect of expiry as distinct from the effect of a song turning 50. However, our results could be invalid if there is a difference in age trends between treated and untreated groups large enough to explain the change in the outcome variable at the time of copyright expiry. For example, if songs released before 1963 saw increasing demand as they aged while songs released in 1963 or later saw falling demand, we could estimate an increase in releases from entering the public domain which actually reflected the diverging trends in age among these two groups.

We rule out this possibility using several approaches. First, we show that results are robust to restricting the sample to songs recorded before 1963 (Column 1 of Table 12). This approach identifies the effect of copyright expiry solely from a comparison of songs older than 50 years with younger songs, and does not exploit the 2013 term extension. The coefficient on Public Domain is robust to this specification (1.252, with a standard error of 0.034). An alternative approach narrows the sample to songs originally released between 1960-1965. These songs should be more similar in their demand trajectories over time than songs released much earlier or later, so we are essentially "matching" songs to a group of songs that are more similar because they were released at around the same time. There is no apparent pre-trend in this narrowed sample, but the large and significant increase in releases at copyright expiry is nonetheless robust to using this narrowed sample. It is also robust to narrowing the sample further (e.g. to narrower windows around 1963, not presented here but available upon request).

Third, we estimate results from the full sample but allow for the possibility that trends in age are not parallel for songs released before 1963 and songs later, by augmenting our baseline specification with a group-specific linear trend. A similar approach is used by Dobkin et al. (2018) to study the economic effects of hospital admissions, a context in which pre-trends in outcome variables obviously complicate the estimation of event study coefficients. This is referred to as a "parametric event study" by Rambachan and Roth (2020) (p. 11). The augmented regression controls for age, year and song fixed effects, as in the specification in Column 3 of Table 4, and adds a linear trend in age interacted with a dummy for the group of songs with original release dates before 1963. This group-specific trend will control for any potential difference in the trends of the two groups. We also present a specification (in Column 6) with the group-specific trend restricted to the 10 years prior to copyright expiry, since the pattern of outcomes immediately preceding copyright expiry is likely to be the most relevant. An additional specification in Column (7) includes interactions between the year of original release and a linear age term, to allow for differences in age trajectories by year of original release. Results are robust to this specification. Column (8) presents results from the US-UK triple difference specification found in Table 4, Column (4), controlling for a differential trend in age for the pre-1963 songs as well as a differential trend in age for songs in the UK market and a differential trend in pre-1963 songs in the UK market. To be precise, we add to the regression in Column (4) an interaction of age with a dummy for pre-1963 cohort, an interaction of age with the UK market dummy, a UK X Pre-63 interaction, and a UK X Pre-63 X Age interaction. The song and age fixed effects control for the main effects of original release year and age. The public domain coefficient is robust to controlling for these additional trend variables.

Following Rambachan and Roth (2020), we also consider the possibility of a kink in the differential trend at the time of copyright expiry. We are unaware of any reason why such a kink would exist. However, Rambachan and Roth (2020) suggest using untreated groups to benchmark potential deviations from parallel trends. One way to obtain a benchmark for a hypothetical change in slope of the differential pre-trend at copyright expiry is to check whether there is a kink in the trend for US releases, which are not affected by UK copyright laws. We thus regressed US releases on fixed effects for age, song and year, as well as the differential age trend (age interacted with a dummy for pre-1963 release) up to age 50 and a separate differential age trend above age 50. The coefficient on this differential age trend is 0.0077 (standard error 0.0015) below the age 50 cutoff and 0.0097 (standard error 0.0016) above the cutoff. Because the slopes before and after the cutoff are very similar in magnitude, with overlapping 95% confidence intervals, this casts substantial doubt on the idea that there could be an increase in the slope that explains our large estimated increase in releases post copyright expiry.[58]

Another approach to estimating a potential kink uses the pre-period coefficients of the event-study regression in Table 4. Rambachan and Roth (2020) provide guidelines for sensitivity analysis that bounds the extent to which the slope of a hypothetical differential trend could change between consecutive periods. For a given M, the amount by which the slope of a differential trend could change from one period to the next, they suggest, "one could...construct an upper bound on the largest change in slope in the pre-period for the groups used in the main event-study of interest. One could then benchmark M in terms of multiples of the largest value observed in the pre-period" (p. 35). In Column 8 of Table 4, the coefficients on the years up to year T=-1 are not significantly different from zero, but the coefficient on T=-1 is negative and significant, with a 95% confidence interval of (-0.271, -0.013). This implies a dip in releases in year T=-1, followed by an increase in year T=0. If we extrapolate the rise in releases from T=-1 to T=0 forward linearly to T=+1, in year T=+1 we would expect to observe releases in the range (0.013, 0.271), in the following year it is 2 times this (0.025,0.542), in the third year (0.038, 0.814) and so on. Even with this extremely conservative assumption, the estimated increase in re-releases upon copyright expiry is outside these bounds for the extrapolated pre-trend.

Turning to the results on performances, we do not observe a pre-trend in Figure 5 that could explain a decline in performances after expiry, but nonetheless we perform a similar robustness analysis to that described above. Results in the set list sample are also robust to restricting the sample to songs released before 1963 (Appendix Table 13 Column 1), or within a narrower range of original release years (Columns 3-4), as well as to augmenting the model to control for group-

[58]The differential trend coefficient of 0.02 in Column (5) of Table 12 would have to increase by 63.5 times one year post expiry, or 15.9 times four years post expiry, to explain the public domain coefficient of 1.271.

specific trends (Columns 5-7). The coefficients in columns (5) and (6) of Table 13 are negative but small. However, since the coefficients in Columns (5) and (6) have quite large standard errors, an intensification in the negative slope of the differential trend of -0.038 in Column (5) could reach the upper limit of the 95% confidence interval by one year post-expiry if the slope decreased by a factor of 17.19, or by 4 years post-expiry if it decreased by a factor of 4.30. To obtain a benchmark for a potential change in slope of the differential pre-trend, we check whether there is a kink in the trend for US performances. We regressed US performances on fixed effects for age, song and year, as well as the differential age trend (age interacted with a dummy for pre-1963 release) below age 50 and a separate differential age trend at age 50 and above. The coefficient on this differential age trend is -0.017 (standard error 0.011) below the age 50 cutoff and -0.021 (standard error 0.012) above the cutoff. The difference in these coefficients is not statistically significant, which itself casts doubt on the idea of a kink in the differential trend. Controlling more carefully for differences in age trajectories across song vintages (with age X origin year interactions in Column 7) leads to an increase in the magnitude of the negative coefficient on Public Domain, implying an 84% reduction in performances.

Column (8) presents results from the US-UK triple difference specification found in Table 5, Column (4), controlling for a differential trend in age for the pre-1963 songs, a differential trend in age for songs in the UK market, an interaction between the pre-1963 and the UK market dummies, as well as a pre-1963 X UK X age interaction.[59] With these interactions in Column (8), effect sizes are similar and inference remains consistent with Column (4) of Table 5.

In Figure 9, there is a drop in availability on Amazon.co.uk for songs first released in 1963 or later, and one may be concerned with a declining trend in availability from 1959 to 1967. The Amazon event study in Figure 6, for example, does not have an immediately apparent pre-trend except for the negative 1966 coefficient being significantly different from the excluded reference category. To correct for a potential pre-trend, we conduct two robustness tests. First, we include in the regression an interaction between a pre-1963 release indicator with a linear age term. This reduces the public domain coefficient in Table 7 from 1.606 (s.e. 0.049) to 1.259 (s.e. 0.087). Alternatively, to remove outliers observed in the event study we drop all songs that were originally released in 1966, which trivially attenuates the estimated coefficient from 1.606 (s.e. 0.049) to 1.581 (s.e. 0.087).

6 The Saturday Club Data

To explore whether the DSP availability of more obscure artists benefits from public domain status, this section analyzes data on artists who appeared on *The Saturday Club* in the late fifties and sixties. This was one of two popular music programs that aired on the BBC in the late fifties and early sixties.[60] There are 1,079 performers who appeared on these lists between 1958 and 1969.

[59]To be precise, we add to the regression in Column (4) an interaction of age with a dummy for pre-1963 cohort, an interaction of age with the UK market dummy, an interaction of the UK market dummy with a dummy for pre-1963 cohort, and an interaction between age, the UK market dummy, and a dummy for pre-1963 cohort.

[60]We extracted performers names from episode lists posted at http://epguides.com/SaturdayClub/.

We drop 81 performers who appeared on *The Saturday Club* and are also found in our regression sample described above. After matching the data to MusicBrainz to recover original release years, the final sample resulted in 22,831 total songs found on Spotify in any market worldwide by 171 artists that do not appear in our original data set. Note that the song-level sample is conditional on having an exact match by artist name and standardized song name, as well as the song being available in at least one Spotify market globally.

Estimates for the effect of public domain status in this sample are striking, but deserve attention by future researchers due to known empirical challenges in the data. Figure 10 displays the mean availability of songs by year of original release, from 1959 to 1966, split by country. While it appears that songs in the UK may have slightly higher availability in the pre-1963 period (recordings in the public domain) compared to '63-66 period, it is problematic that availability in the US is low in 1959 and appears to systematically grow over year of original release through 1966. This may be due to the fact that the popularity of the BBC's *Saturday Club* grew over time, and more popular artists appeared in later years of the show. Thus, while the most obscure *Saturday Club* artists may only have notable demand in the UK, more successful artists appeared later in the show with global appeal, correlating with year of original release.

Despite these empirical challenges, the results are striking. Column 1 of Table 14 compares availability between the US and UK while controlling for artist, country, and age fixed effects, implying a 39% increase in availability for those sound recordings that are in the public domain in the UK. Column 2 interacts original release year with a UK indicator, and while songs released in the pre 1963 period generally have greater availability in the UK of between 18 and 31 percent, there does appear to be a significant pre-trend in the sense that years 64-66 have systematically lower availability than the excluded year of 1963. Again, estimation here is complicated by the fact that US availability for this sample grows over time from 1959 to 1966, despite no break in copyright status. Due to this fact that US availability pre-1963 may be a flawed counterfactual for UK availability pre-1963, we drop US observations from the sample and compare availability in the UK in the 1959-1962 period to the 1963-1966 period, with the maintained assumption that the availability of 1963-1966 songs provides a counterfactual for the availability of pre 1963 songs had those songs not lapsed into the public domain. Column 3 compares the availability of public domain recordings in the UK (1959-1962) to copyright protected recordings (1963-1966), and implies that songs in the public domain in this sample are about 5.1% more likely to be available on Spotify in the UK than copyright protected songs. Column 4 adds a linear control for age and shows a similar effect.

Columns (5) - (6) of Table 14 study availability on Spotify at the artist level. In this analysis, we study the 998 artists that appeared on *The Saturday Club* and are not in our main sample. We record if each artist has any songs on Spotify in the US and UK market. Since these analysis are run at the artist level, we use year of appearance on the Saturday Club as a proxy for original release year. Observations in the UK from artists that appeared in the pre-1963 period are considered in the public domain. Column 5 controls for country and age fixed effects and implies a 4.6% increase in availability for artists who appeared on *The Saturday Club* before 1963 in the UK. Column 6 compares the availability of artists who appeared in the pre 1963 period (1959-1962) to those who appeared after, only within the UK market and shows no significant difference. Altogether, these results are suggestive of a possible effect of public domain status among these more obscure artists, but warrant further research.

7 Probability of touring

Appendix Table 11 presents results collapsed to the artist-year level. Column (1) presents coefficients from a regression of the dummy capturing whether the artist toured in the UK in a given year on the percentage of the artist's songs in the public domain in that year, controlling for artist and year fixed effects and a quadratic in the average age of the artist's songs. We find that a 1 percent increase in the number of songs in the public domain in the UK reduced the probability of touring in the UK by 0.4 percentage points (significant at the 1% level). By contrast, Column 2 shows that there is no significant difference in the probability of touring in the US as the percentage of songs in the public domain in the UK increases (the coefficient estimate is -0.0002 percentage points with a standard error of 0.0009 percentage points). We also examined the total number of songs performed per year as a function of the percentage of an artist's songs in the public domain (Columns 3-4). Once again, there is a statistically significant decline in the number of songs performed per year in the UK as the percentage of songs in the public domain increases (Column 3). There is no significant effect for the US (Column 4).

Dataset Construction

OfficialCharts/Musicbrainz Data

Weekly top 20 album charts were collected from the Official Charts Company (OCC) between 1960 and 1965. Artist and group names were then hand-matched to the unique database identifiers in the MusicBrainz database, with 138 artists successfully matched between the datasets.[61] To collect a sample of songs by more obscure artists for robustness checks in our Spotify analysis, we also collected weekly top 50 singles charts from OCC between 1960 and 1965. Artist and group names were then hand-matched in an identical manner as described above to create a sample of 41 relevant artists, who do not also appear in the albums charts, have only one song in weekly top singles charts, and whose song appears on the charts for two or fewer weeks.[62]

A local Musicbrainz Virtual Machine was then used for generating the data-sets via SQL queries.[63] All recordings and the releases of such recordings by the sample artists were collected - thus we pick

[61] Official Charts tracked just the Top 10 albums during the first two months of 1960.
[62] Official Charts tracked just the Top 30 singles during the first three months of 1960.
[63] https://musicbrainz.org/doc/MusicBrainz_Server/Setup

up not only albums/singles released by the artists, but also compilation albums featuring various artists. Data was also collected on the country of release, the format of release (e.g., CD, SACD, Digital), and the release label.

Song titles were standardized by: a) converting titles to lowercase, b) stripping accent marks, and c) removing punctuation marks. To ensure a reliable year of release, we measured the original year of release as the earliest original year listed between both our MusicBrainz data and data from the Discogs music database (data.discogs.com). Artists in our MusicBrainz data were carefully hand-matched to the corresponding artists in the Discogs database, and song titles in the Discogs data were standardized via the aforementioned method. Further, we manually confirm original release years for any song in the sub-sample used for Amazon analysis (described above), performed after age 40, or re-released after age 40 by an artist with an error rate above 20% within the Amazon sample. Validation of original release years was done using sites such as Wikipedia, Secondhandsongs.com, and 45cat.com

Data are missing on the original release year of the song for 7,831 of 716,685 total observed song releases in the raw data, and 3,476 of 599,624 observations when the sample is restricted to official releases. These observations are dropped from the sample. Country of release is missing for 115,914 of these observations, or 70,244 for official releases. Observations with missing data on country of release and year of release are dropped from the dataset. In previous versions of this manuscript both official releases and "bootleg" releases (bootlegs account for 66,023 total releases) were included in the data set, but the current analysis considers only official releases.[64] Results are robust between including or omitting bootlegs. Data are collapsed to the release-level, such that a given release with multiple variations on the same recording is counted only once.

Set List Data

Musician set list data was collected from setlist.fm using their REST API documented at https://api.setlist.fm/docs/1.0/index.html. Setlist.fm's database tracks artists using the Musicbrainz GID, the same unique identifier used by the MusicBrainz database. Of the artists in our sample, 99 appeared in the Setlist.FM data, as matched by MusicBrainz GIDs. All set lists were collected via the web API for these 99 artists, resulting in 16,847 total concert set lists encompassing 295,232 total song performances. Songs in this data were manually standardized by stripping extraneous characters and standardizing case. They were then matched to equivalently standardized song/artist combinations in the re-release data and matches were kept for analysis.

Spotify Data

Artists in our MusicBrainz data were hand matched to Spotify's artist unique identifiers (URIs). Catalog information, including geographic availability, was then downloaded from Spotify's API

[64] MusicBrainz also includes a small number of promotional releases in addition to official and bootleg releases.

for all of the artists in our MusicBrainz dataset, see https://developer.spotify.com/web-api/get-track/ for fields obtained. Song titles in the Spotify dataset were matched to the MusicBrainz data by artist. We standardized the Spotify song titles to match our standardized MusicBrainz titles: titles were converted to lowercase, punctuation was stripped, and the word remaster was stripped.[65] Remaining unmatched titles between our Spotify and MusicBrainz data were then manually matched in order to properly match titles with alternative spellings.

Amazon Music Data

To obtain data on the availability of tracks as MP3 downloads on Amazon Music, we took a random sample of songs from the Musicbrainz database, and searched the digital music pages on Amazon.co.uk and Amazon.com for MP3 downloads matching the song title and artist name. The random sample consisted of 2025 songs, stratified by original release year between years 1959 to 1967 (225 songs per year) from our dataset of physical releases. The original release years for all songs in this sample were checked by hand to ensure the precision of results. We crawled the search results pages and ran python code that extracted the number of search results listed in response to structured search queries for song title and artist name. We also recorded song title and artist names on each page of search results.

Given that a naive count of these search results may inflate availability due to similar terms and cover songs by similar artists, we structured our search queries using URL parameters to ensure that: 1) only songs from the Digital Music store were returned using the appropriate URL parameter, excluding search results for albums (which would lead to double-counting when an album and song share a title) and physical editions, 2) results were restricted to Amazons unique indexed artist name for the focal song using the p_lbr_music_artists_browse-bin URL parameter. Each page of search results was scraped and song titles were then matched with our sample to count the total number of available editions on both the US and UK store. Although we believe we have eliminated all cases in which the search result count was substantially overstated, our data may in some cases slightly misstate the number of versions of a song available due to the nature of Amazon's algorithm for displaying search results, and the difficulty of obtaining a precise count of the number of versions of an MP3 download available on the site. However, because we compare the number of search results for the same song on Amazon's US and UK sites, any overcounting due to Amazon's search algorithm should be differenced out in this comparison. Any remaining errors should merely introduce noise into our sample and should not bias the results.

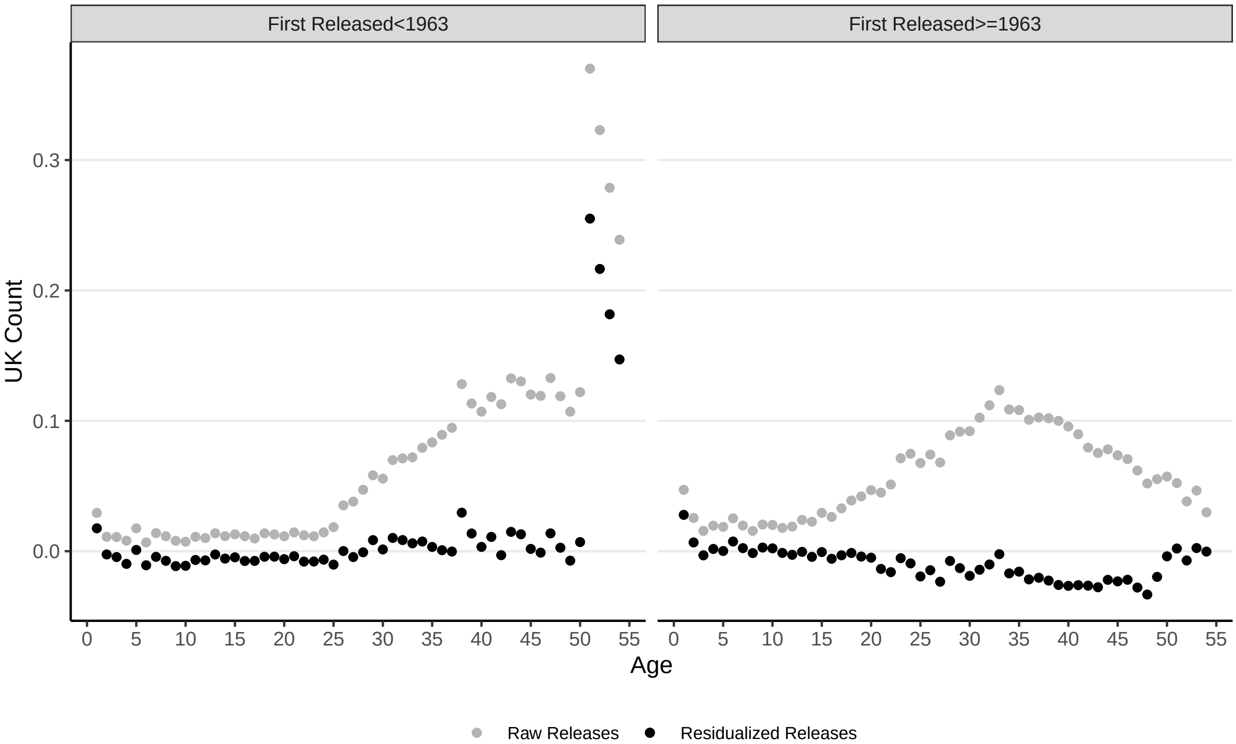
We also use this random sample, stratified by year, to validate our Spotify results. Overall, 79.0% of songs in our Spotify dataset are available in the US or UK on Spotify. To ensure this availability rate is not driven by any errors in our data matching process, we look at this sample of songs that had been mechanically matched to Amazon and were found at a rate of around 78.9% in February

[65] Copyrights for remastered sound recordings cover only those elements of the new fixation that differ from the original. See guidance from PPL in the UK, http://www.ppluk.com/Documents/Distribution/Guidance%20to%20PPL%20Members%20on%20Remasters.pdf

2020. We re-ran our Spotify analysis. The results were consistent, thus we have no reason to believe any tracks that were imperfectly matched would have any influence on our findings.

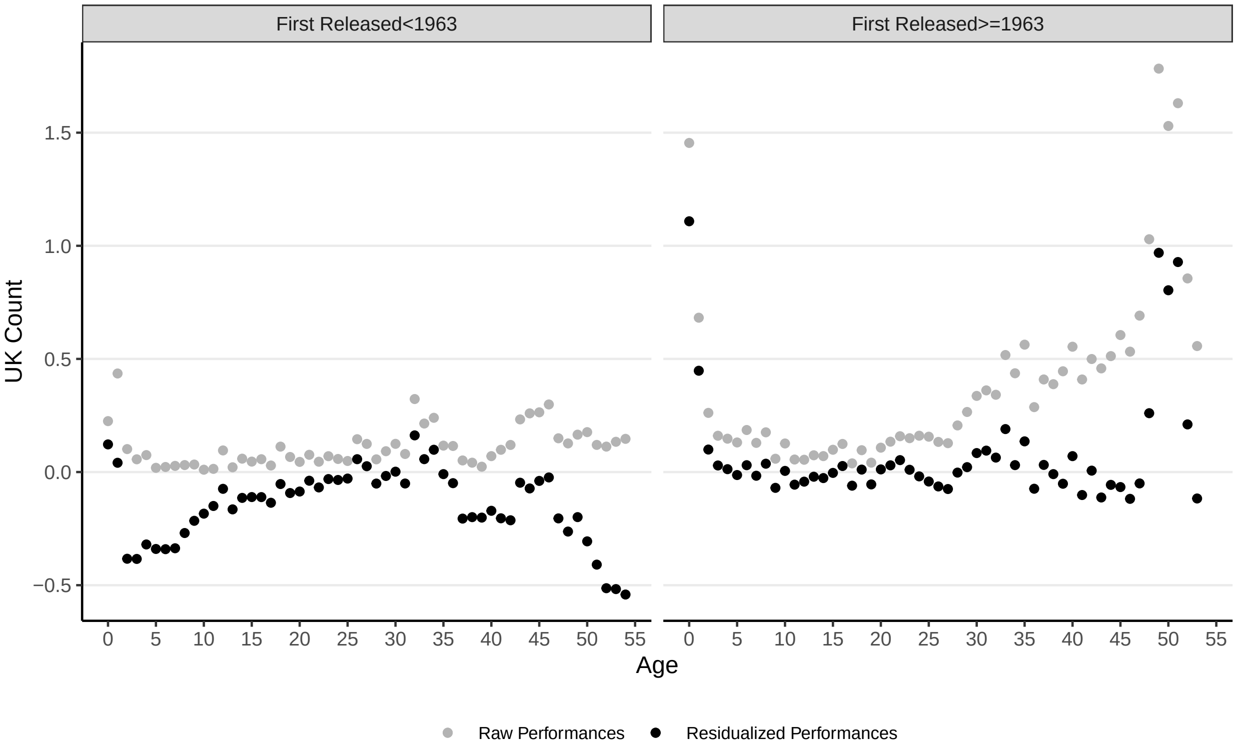
8 Results Appendix

Figure 7: Raw Effect of Copyright on Physical Releases in the UK



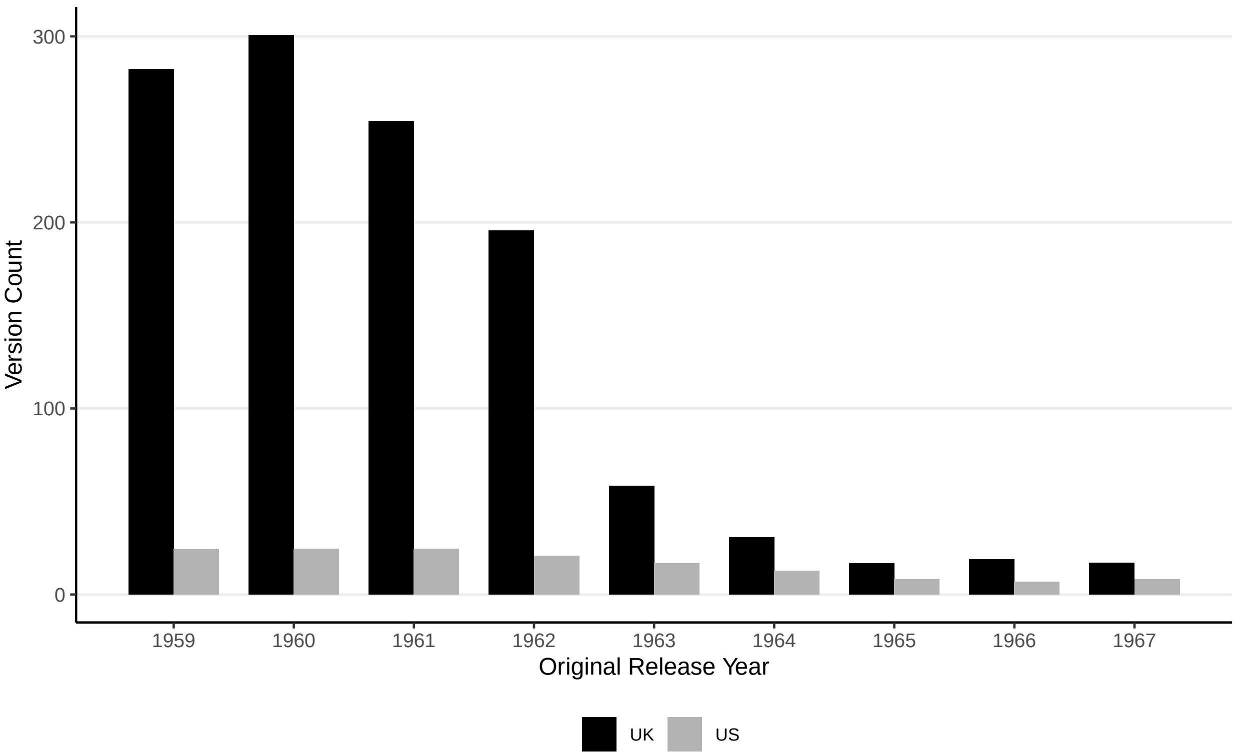
Notes: This figure shows the average number of physical releases in the UK by song age. Residualized mean releases are computed from the residuals of a regression of the outcome variable on dummies for year t. The left panel is restricted to songs released before 1963, for which the recording copyright term remained fixed at 50 years. The right panel is restricted to songs released in 1963 or later, which had copyright terms extended at the end of 2013. The estimation sample includes all observations in which age is less than 55 years.

Figure 8: Raw Effect of Copyright on Concert Performances in the UK



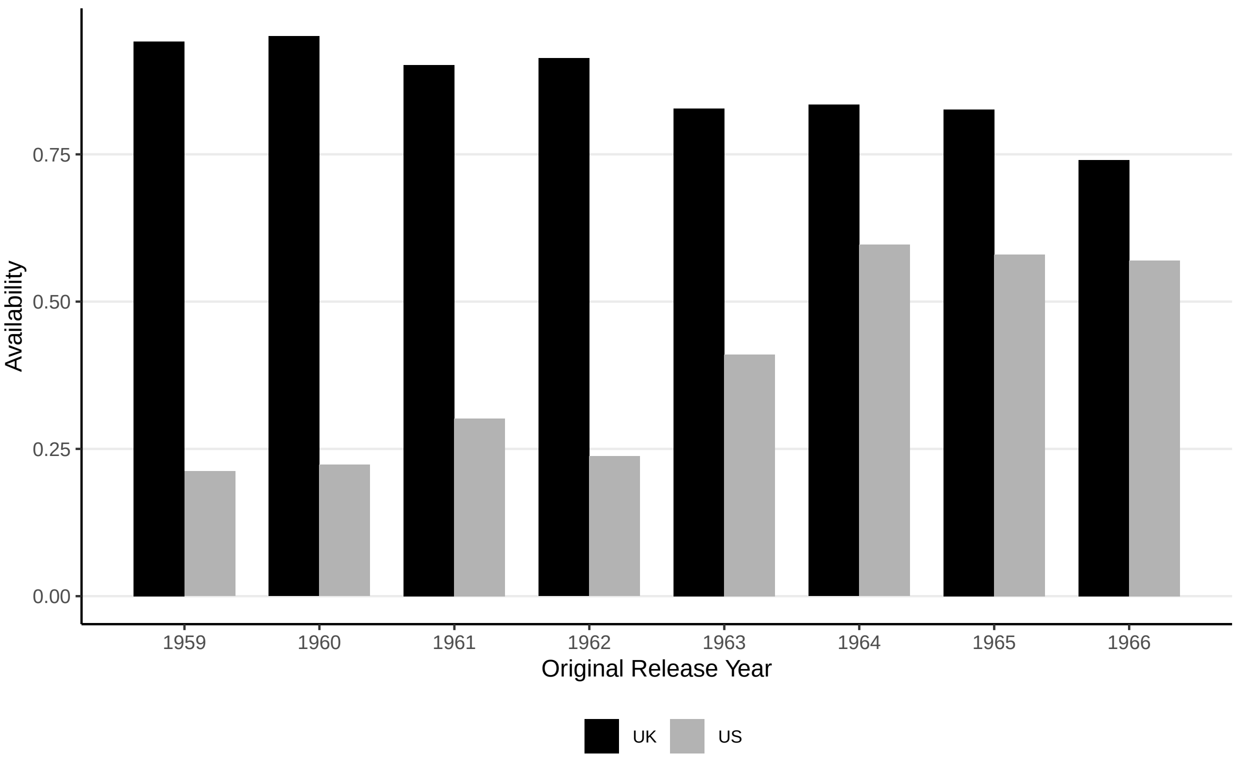
Notes: This figure shows the average number of times songs were performed in the UK by song age. Residualized mean releases are computed from the residuals of a regression of the outcome variable on dummies for year t. The left panel is restricted to songs released before 1963, for which the recording copyright term remained fixed at 50 years. The right panel is restricted to songs released in 1963 or later, which had copyright terms extended at the end of 2013. The estimation sample includes all observations in which age is less than 55 years.

Figure 9: Raw Effect of Copyright on Amazon Versions



Notes: This figure shows the average number of versions for songs, by original release year, on the UK and US Amazon platforms.

Figure 10: Percentage of *Saturday Club* Songs on Spotify, by Market



Notes: This figure illustrates the availability of songs in our song-level *Saturday Club* sample by year of original release and country. Availability measures the mean availability on Spotify UK and Spotify US in 2021.

Table 8: Robustness of Results on Releases

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	DropTop5	DropTop5	FullSample	FullSample	Linear ln(Releases)	Linear ln(Releases)	US Placebo
Public Domain	1.167***	1.236***	1.359***	1.146***	0.129***	0.133***	0.035
	(0.168)	(0.083)	(0.182)	(0.060)	(0.020)	(0.003)	(0.097)
Age FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Artist FE	Yes	No	Yes	No	Yes	No	No
Song FE	No	Yes	No	Yes	No	Yes	Yes
N	811,109	484,553	966,550	634,906	922,182	922,182	252,598

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table displays estimated coefficients from a Poisson regression. In columns (1)-(4) the dependent variable is the number of UK releases of song i in year t. All columns include year and age fixed effects. Columns (1) and (2) exclude the top 5 artists in the sample in terms of UK releases (The Kinks, Elvis Presley, The Beatles, Frank Sinatra, and Jerry Lee Lewis). Columns (3) and (4) include songs that are more than 54 years old. Columns (5) and (6) use a linear specification, where the dependent variable is the log of (UK releases +1). Column (7) is a placebo regression with the dependent variable being number of US releases of song i in year t, and includes songs released between 1960-1965. Age fixed effects control for the number of years since original release. Year fixed effects control for the year of release. Song fixed effects control for the artist-song combination. Standard errors are clustered at the level of the respective group fixed effect, artist or song.

Table 9: Robustness of Results on Setlists

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	ln(Performances)	ln(Performances)	FullSample	FullSample	TouringYrs	TouringYrs	DropT5	DropT5	US Count
Public Domain	-0.142*	-0.140***	-1.802***	-1.913***	-1.522***	-1.537***	-1.788***	-1.898***	-0.146
	(0.077)	(0.035)	(0.602)	(0.365)	(0.458)	(0.330)	(0.608)	(0.364)	(0.292)
Age FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Artist FE	Yes	No	Yes	No	Yes	No	Yes	No	No
Song FE	No	Yes	No	Yes	No	Yes	No	Yes	Yes
N	109,583	109,575	98,402	56,835	66,168	37,005	97,804	56,266	30,182

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Columns (1) and (2) are linear models that use the log of (UK performances+1) as the dependent variable. Columns (3) through (9) are Poisson models. Columns (3) and (4) use the full sample of performances without restricting the sample to songs that are less than 55 years old. Columns (5) and (6) only include years in which the artists has at least one live performance. Columns (7) and (8) drop the top five artists by total performances: The Beach Boys, Bob Dylan, Frank Sinatra, The Who, and The Rolling Stones. Column (9) uses US performances as the dependent variable, and includes songs released between 1960-1965. Age fixed effects control for the number of years since original release. Year fixed effects control for the year of release. Song fixed effects control for the artist-song combination. Standard errors are clustered at the level of the respective group fixed effect, artist or song.

Table 10: Robustness of Results on Digital Streaming

	(1) Naive	(2) Full Sample	(3) Full Sample	(4) Drop Unavailable	(5) Drop T5	(6) Amazon Sample	(7) Age Restricted
Public Domain (pre1963 x UK)	0.014	0.017	0.017	0.011	0.027	-0.017	0.027
	(0.016)	(0.015)	(0.015)	(0.020)	(0.018)	(0.023)	(0.019)
Age FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Artist FE	No	Yes	No	Yes	Yes	Yes	Yes
Song FE	Yes	No	Yes	No	No	No	No
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	37,032	39,718	39,718	29,244	31,636	1,934	32,360

Standard errors in parentheses

[*] $p < 0.10$, [**] $p < 0.05$, [***] $p < 0.01$

Notes: This table displays coefficients from a linear probability model. In all specifications, the dependent variable is a binary variable that equals 1 if song i is available in geographic market k on the digital music streaming platform Spotify as of September 2017. The estimation sample is restricted to songs with an original release year before 1975. All specifications control for country fixed effects. Column (1) also controls for song fixed effects. Columns (2) and (3) include the sample of obscure tracks from the singles charts. Column (4) restricts the sample to songs that are available on Spotify either in the US or the UK. Column (5) drops the top five artists in our Spotify data: Frank Sinatra, Elvis Presley, Ella Fitzgerald, Peggy Lee, and Sarah Vaughan. Column (6) restricts the population to songs that were analyzed in our random Amazon sub-sample. Column (7) drops songs older than 60 years. Age fixed effects control for the number of years since original release. Song fixed effects control for the artist-song combination. Country fixed effects control for the US market versus the UK market. Standard errors are clustered by artist.

Table 11: Analysis of touring probability and total performance counts

	(1) UK Tour	(2) US Tour	(3) UK Count	(4) US Count
Percent Songs in Public Domain	-0.004***	-0.000	-0.084***	-0.018
	(0.001)	(0.001)	(0.025)	(0.023)
Avg(Age)	-0.014	-0.042***	-0.485**	-0.363***
	(0.008)	(0.014)	(0.214)	(0.104)
Avg(Age)^2	0.000	0.000	0.004	0.001
	(0.000)	(0.000)	(0.003)	(0.001)
Year FE	Yes	Yes	Yes	Yes
Artist FE	Yes	Yes	Yes	Yes
N	3,294	3,294	2,262	2,618

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: The unit of observation is an artist-year. The dependent variable in Column (1) is a dummy equal to 1 if artist j toured in the UK in year t. The dependent variable in Column (2) is a dummy equal to 1 if artist j toured in the US in year t. The dependent variable in Column (3) is the count of total performances by artist j in the UK in year t. The dependent variable in Column (4) is the total count of performances by artist j in the US in year t. Columns (1-2) are estimated using Ordinary Least Squares, Columns (3-4) using fixed effects Poisson. "Avg(Age)" is calculated as the mean age by artist-year. Standard errors are clustered by artist.

Table 12: Robustness of Releases Results to Trend Assumptions

	(1) Released < 1963	(2) Released < 1963	(3) 1960-65	(4) 1960-65	(5) Full Sample	(6) Full Sample	(7) Full Sample	(8) UK vs US
Public Domain	1.252***				1.271***	1.390***	1.004***	0.992***
	(0.034)				(0.066)	(0.065)	(0.027)	(0.029)
T = -4		-0.398***	-0.280	-0.039				
		(0.129)	(0.210)	(0.098)				
T = -3		-0.203	-0.315	-0.038				
		(0.130)	(0.203)	(0.095)				
T = -2		-0.202	-0.260	-0.111				
		(0.135)	(0.196)	(0.099)				
T = -1		-0.212***	-0.138	-0.095				
		(0.065)	(0.142)	(0.096)				
T = 0		0.000	0.000	0.000				
		(.)	(.)	(.)				
T = 1		1.189***	1.386***	1.584***				
		(0.098)	(0.179)	(0.086)				
T = 2		1.238***	1.166***	1.434***				
		(0.096)	(0.220)	(0.096)				
T = 3		1.288***	0.925***	1.279***				
		(0.101)	(0.317)	(0.119)				
T = 4		1.290***	0.386	0.861***				
		(0.095)	(0.295)	(0.132)				
Age × Treated					0.020***	0.005***		
					(0.002)	(0.001)		
Age × UK								0.000
								(0.001)
UK × pre-63								-0.617***
								(0.053)
Age × UK × pre-63								0.008***
								(0.001)
Age FE	No	No	No	No	Yes	Yes	No	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Artist FE	No	Yes	Yes	No	No	No	No	No
Song FE	Yes	No	No	Yes	Yes	Yes	Yes	Yes
Orig. Year FE	No	No	No	No	No	No	Yes	No
N	242,734	354,638	358,470	227,563	572,665	572,665	572,665	1,555,256

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Releases results, robustness to differential trends. Columns (1) and (2) restrict the sample to songs released before 1963. Columns (3) and (4) restrict the sample to songs released between 1960 and 1965. Column (5) controls for age, year and song effects, as well as a linear term for age interacted with a indicator variable for original release year before 1963. Column (6) duplicates Column (5), except the linear trend is restricted to age 40 and up. Column (7) includes fixed effects for year, song and year of original release, as well as year of original release interacted with a linear age term. Column (8) is based on the UK versus US sample with controls for differential trends in age for the pre-1963 releases, UK releases, and pre-1963 releases in the UK.

Table 13: Robustness of Setlist Results to Trend Assumptions

	(1) Released < 1963	(2) Released < 1963	(3) 1960-65	(4) 1960-65	(5) Full Sample	(6) Full Sample	(7) Full Sample	(8) UK vs US
Public Domain	-0.987**				-1.347***	-1.260***	-1.834***	-0.874**
	(0.462)				(0.354)	(0.366)	(0.288)	(0.358)
T = -4		0.269	1.856***	1.841***				
		(0.264)	(0.632)	(0.543)				
T = -3		-0.297	1.385**	1.239**				
		(0.588)	(0.568)	(0.525)				
T = -2		-0.374*	1.779**	1.976***				
		(0.213)	(0.783)	(0.554)				
T = -1		-0.104	2.589***	2.781***				
		(0.338)	(0.750)	(0.544)				
T = 0		0.000	0.000	0.000				
		(.)	(.)	(.)				
T = 1		-0.752**	0.043	-1.167				
		(0.317)	(0.907)	(0.947)				
T = 2		-1.120***	-0.528	-1.813**				
		(0.281)	(1.263)	(0.780)				
T = 3		-1.934***	-0.209	-1.461				
		(0.366)	(1.424)	(1.049)				
T = 4		-1.856***	-0.008	-1.066				
		(0.514)	(1.225)	(0.937)				
Age × Treated					-0.038***	-0.026***		
					(0.009)	(0.005)		
Age × UK								0.007
								(0.004)
UK × pre-63								-1.011***
								(0.362)
Age × UK × pre-63								0.018
								(0.014)
Age FE	No	No	No	No	Yes	Yes	No	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Artist FE	No	Yes	Yes	No	No	No	No	No
Song FE	Yes	No	No	Yes	Yes	Yes	Yes	Yes
Orig. Year FE	No	No	No	No	No	No	Yes	No
N	9,917	24,399	36,641	22,653	56,266	56,266	56,195	193,568

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Setlist results, robustness to differential trends. Columns (1) and (2) restrict the sample to songs released before 1963. Columns (3) and (4) restrict the sample to original release years 1960-65. Column (5) controls for age, year and song effects, as well as a linear term for age interacted with a dummy for release year before 1963. Column (6) is the same as Column (5), only the trend is restricted to age 40 and up. Column (7) controls for fixed effects for year, song and year of original release, as well as year of original release interacted with a linear age term. Column (8) is based on the UK versus US sample with controls for differential trends in age for the pre-1963 releases, UK releases, and pre-1963 releases in the UK.

Table 14: Saturday Club Artists on Spotify

	(1) US vs UK	(2) US vs UK	(3) UK pre/post	(4) UK pre/post	(5) US vs UK	(6) UK pre/post
Public Domain	0.389***		0.051***	0.052**	0.046***	-0.020
	(0.049)		(0.016)	(0.021)	(0.016)	(0.033)
UK					0.029***	
					(0.008)	
UK x Release = 1959		0.312***				
		(0.091)				
UK x Release = 1960		0.309***				
		(0.066)				
UK x Release = 1961		0.182*				
		(0.092)				
UK x Release = 1962		0.258***				
		(0.067)				
UK x Release = 1963		0.000				
		(.)				
UK x Release = 1964		-0.180***				
		(0.059)				
UK x Release = 1965		-0.171***				
		(0.065)				
UK x Release = 1966		-0.247***				
		(0.065)				
Age				-0.000		
				(0.005)		
Artist FE	Yes	Yes	Yes	Yes	No	No
Country FE	Yes	Yes	No	No	Yes	No
Orig. Year FE	Yes	Yes	No	No	Yes	No
N	45,662	45,662	22,821	22,821	1,996	998

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Columns (1) and (2) compare availability between the US and UK at the song-level for songs by Saturday Club artists, with dependent variable = 1 if song i is available in market k (UK or US) in 2021. Columns (1) and (2) include artist, country, and original release year fixed effects. Column (2) introduces an interaction term of a UK indicator with year of original release (1963 omitted). Columns (3) and (4) limit observations to UK-only, and compare availability of songs released before 1963 (public domain in the UK) to those released during or after 1963 (copyright protected). Column (4) includes a linear age control. Columns (5) and (6) use the artist-level sample, where the dependent variable = 1 if any song by the *Saturday Club* artist is on Spotify in market k (UK or US) in 2021 and the original release year is proxied by the artist's average year of appearance on the *Saturday Club*. Column (5) includes all artists and controls for country and release year fixed effects. Columns (6) limits observations to the UK-market only, and compares the availability of artists on Spotify UK before 1963 to those with average year of release equal to or greater than 1963.

Table 15: Panel data collapsed into five-year age bins

	(1) Releases	(2) Releases	(3) Releases	(4) Performances	(5) Performances	(6) Performances
Public Domain	1.369***	1.302***		-1.824***	-2.199***	
	(0.118)	(0.027)		(0.586)	(0.321)	
Pre'63 x Ages [31,35]			0.202*			1.019*
			(0.106)			(0.527)
Pre'63 x Ages [36,40]			0.261**			-0.481
			(0.109)			(0.508)
Pre'63 x Ages [41,45]			0.170			0.395
			(0.104)			(0.441)
Pre'63 x Ages [46,50]			0.000			0.000
			(.)			(.)
Pre'63 x Ages [51,55]			1.167***			-1.427***
			(0.108)			(0.461)
Age	0.069***	0.024***	0.066***	0.054**	0.024	0.043*
	(0.009)	(0.009)	(0.010)	(0.026)	(0.030)	(0.026)
5-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Artist FE	Yes	No	Yes	Yes	No	Yes
Song FE	No	Yes	No	No	Yes	No
N	196,560	123,483	196,560	21,278	11,108	21,278

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: The panel data on releases and performances for a given song is collapsed into five-year bins, e.g., [1960,1965), [1965,1970), etc. In these regressions, the dependent variable is the count of the outcome variable in a given five-year bin, the main variable of interest is mean of public domain dummy, and control variables are the 5-year binned means of song age a within a year bin, 5-year bin fixed effects, and song or artist fixed effects. Columns (1) and (4) controls for five-year and artist fixed effects and column (2) and (5) controls for five-year and song fixed effects. Column (3) and (6) interact an indicator for songs released before 1963 with indicators for the binned average song age within the bin. Songs released during or after 1963 are pooled with the excluded category of Age = [46,50]. Indicators for ages younger than [31,35] are included in the estimation but omitted from the table. Standard errors are clustered at the level of the respective group fixed effect, artist or song.