Information Systems Research

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Video Killed the Radio Star? Online Music Videos and Recorded Music Sales

Tobias Kretschmer, Christian Peukert

informs
http://pubsonline.informs.org/journal/isre

INFORMATION SYSTEMS RESEARCH
Vol. 31, No. 3, September 2020, pp. 776–800
ISSN 1047-7047 (print), ISSN 1526-5536 (online)

Video Killed the Radio Star? Online Music Videos and Recorded Music Sales

Tobias Kretschmer,[a,b] Christian Peukert[c,d]

[a] Ludwig-Maximilians-University of Munich, 80539 Munich, Germany; [b] Centre for Economic Policy Research, London EC1V 0DX, United Kingdom; [c] Católica Lisbon School of Business and Economics, Universidade Católica Portuguesa, 1649-023 Lisbon, Portugal [d] Center for Law and Economics, ETH Zurich, 8092 Zurich, Switzerland

Contact: t.kretschmer@lmu.de, https://orcid.org/0000-0002-7830-7825 (TK); christian.peukert@ucp.pt, https://orcid.org/0000-0003-3997-8850 (CP)

Received: February 7, 2018
Revised: December 14, 2018; July 13, 2019
Accepted: September 24, 2019
Published Online in Articles in Advance: June 10, 2020

https://doi.org/10.1287/isre.2019.0915

Abstract. We study the heterogeneous effects of online video platforms on the sales volume and sales distribution of recorded music. Identification comes from two natural experiments in Germany. In 2009, virtually all music videos were blocked from YouTube as a result of a legal dispute. In 2013, the dedicated platform Vevo entered the market, making videos of a large number of artists available overnight. Our estimates suggest that restricting (enabling) access to online videos decreases (increases) recorded music sales on average by about 5%–10%. We show that the effect operates independently of the nature of video content, suggesting that user-generated content is as effective as official content. Moreover, we highlight heterogeneity in this effect: online music videos disproportionally benefit sales of new artists and sales of mainstream music.

History: Param Singh, Senior Editor; Gordon Burtch, Associate Editor.

Funding: This work was supported by FCT-Portuguese Foundation of Science and Technology [Grants UID/GES/00407/2013 and FCT-PTDC/EGE-OGE/27968/2017].
Supplemental Material: The online appendix is available at https://doi.org/10.1287/isre.2019.0915.

Keywords: digital distribution platforms • user-generated content • natural experiment

1. Introduction

Digitization has brought important changes to the recorded music industry. Thanks to a substantial decline in the fixed cost of production, distribution, and promotion, the number of new songs on the market has tripled (Aguiar and Waldfogel 2016, 2018a). However, the precise way that digital distribution and promotion affect artists' revenue streams (with implications for incentives to invest in new content in the long run) is not fully understood. We compare *open* and *closed* content distribution platforms in their effects on revenues in other distribution channels and highlight important heterogeneity in those effects. Digital platforms distributing content differ in the amount of control rights holders have over available content and the corresponding revenue model. The spectrum ranges from *unlicensed platforms*, such as (the original) Napster or Megaupload, to *licensed open platforms*, such as YouTube and SoundCloud that allow anyone to upload content, to *licensed closed platforms*, such as iTunes and Spotify, where only the rights holders themselves can contribute.

By definition, unlicensed platforms do not compensate rights holders. Indeed, the relatively large literature on digital piracy suggests that unlicensed consumption mostly substitutes for licensed consumption.[1] Licensed platforms typically pay some share of sales, advertising revenues, or per-use royalties. An emerging literature has started to look into closed platform streaming services and typically finds that services such as Spotify substitute for piracy and digital ownership but leave aggregate revenues largely unchanged (e.g., Thomes 2013, Nguyen et al. 2014, Wlömert and Papies 2016, Aguiar and Waldfogel 2018b, Datta et al. 2018). Surprisingly little is known about licensed open platforms. They are particularly interesting because they create a market for derivative works. Rights holders are compensated for third-party use of their work without requiring individual licensing contracts. For example, on YouTube, where user-generated content (UGC) comprises the majority of available content (Liikkanen and Salovaara 2015), rights holders can get a share of advertising revenues generated from third-party content.

We study how the open platform YouTube affects sales of recorded music and compare its effect with that of Vevo, a closed platform. We then investigate how availability on YouTube affects the type and variety of music consumers demand on other channels (Piolatto and Schuett 2012, Datta et al. 2018). As with many other digital platforms, YouTube offers tools that allow for search and social interaction and provides up-to-date lists of the most popular videos and automated recommendations to help consumers comb through the vast amount of content available on the platform (Zhou et al. 2016), although its role in shaping the popularity distribution is not clear ex ante (Fleder and Hosanagar 2009, Tucker and Zhang 2011, Oestreicher-Singer and Sundararajan 2012b, Susarla et al. 2012, Godinho de Matos et al. 2016).

A unique setting in the German market helps us establish a causal link between YouTube availability and other consumption channels for recorded music. As a result of a royalty dispute between YouTube and the de facto monopolist royalty collection society that represents artists and publishers (not record labels), the Society for Musical Performing and Mechanical Reproduction Rights (Gesellschaft für musikalische Aufführungs- und mechanische Vervielfältigungsrechte (GEMA)), YouTube blocked access to almost all videos containing music in Germany on April 1, 2009.[2] For example, 85% of the 689 music videos in the list of the 1,000 most viewed videos globally were blocked in Germany, whereas the same content remained accessible in a vast majority of other countries. For instance, 99% of those videos could be accessed in the United States.[3] This standoff persisted until a consortium of record labels negotiated its own deal with GEMA and launched the dedicated platform Vevo on October, 1, 2013, which in most other countries is simply a channel on YouTube.

We use scanner data on the weekly numbers of units sold on physical media and as digital downloads for the 1,000 highest-grossing songs on the German market and match them with rich meta-information at the release and artist levels. For 2013, we also observe the weekly number of (free and paid) streams per song. For each song in our data, we collect information on whether a corresponding video was available on YouTube (or Vevo). We estimate a difference-in-differences model to compare sales of songs *with* videos with sales of songs *without* videos four weeks before and four weeks after the natural experiment(s) we observe.

Our results provide strong evidence that YouTube is complementary to other licensed consumption channels, at least in the short run. Across a variety of different specifications, our most conservative estimates suggest that removing access to music videos on YouTube *reduces* total weekly sales by approximately 6% on average. This result is robust to a number of falsification exercises, including placebo tests and data from Austria, a country that shares language and cultural history with Germany but was not affected by the blocking on YouTube. We show that the size of our estimated effect does not vary across songs that have a higher share of user-generated videos on YouTube, suggesting that UGC has a complementary effect very similar to that of official content. Strikingly, weekly sales and streams *increase* by a similar amount when Vevo, a closed platform without UGC, enters the market four years later. We do not find evidence that suggests that the complementary effect is moderated by overall popularity as measured in sales. We then study heterogeneity across songs and find that YouTube disproportionately benefits sales of new artists and mass-market artists (mainstream music). Finally, we show that the experiments we study in this paper affect the composition of not only sales/streams of songs but also aggregated sales/streams of songs. However, our data preclude us from studying the effect of online music videos on bundled sales of songs, for example, through albums. We discuss these limitations in the context of the related literature.

By studying the (un)availability of a licensed open platform for music consumption and its heterogeneous effect on demand on other consumption channels, we make several contributions. Specifically, we emphasize the effect of digital distribution platforms on product discovery rather than just their substitution of paid channels. Demonstrating that mainstream artists and new artists benefit disproportionally from digital distribution platforms also adds to the literature on the effect of digitization on the sales distribution ("long tail"; Brynjolfsson et al. 2010, 2011). Furthermore, our results may be informative in the context of cultural trade policy (Hervas-Drane and Noam 2017) and the debate on the reform of the compulsory licensing rules of interactive digital services (Lenard and White 2015).

2. Related Work and Research Framework

2.1. Channel Competition: Displacement or Promotion?

A growing literature seems to have established that unlicensed consumption (digital piracy) harms sales of licensed products (e.g., Bhattacharjee et al. 2007, Adermon and Liang 2014, Danaher et al. 2014), with substantial variation in estimated displacement rates (Hui and Png 2003; Rob and Waldfogel 2006, 2007; Zentner 2006; Oberholzer-Gee and Strumpf 2007; Liebowitz 2016). Recent empirical evidence on digital piracy (Peukert et al. 2017) supports the theoretical work, arguing that unpaid consumption can increase sales in licensed channels because of demand externalities (Takeyama 1994) or because consumers can sample (vertical or horizontal) product quality a zero marginal cost (Peitz and Waelbroeck 2006). The growing adoption of licensed services, technological advances, and business model innovations has changed how and at what cost consumers access digital goods. Theoretically, (low-priced) licensed online offerings can combat piracy (Thomes 2013), which is empirically supported in several settings (Danaher et al. 2010, Papies et al. 2011, Poort and Weda 2015, Aguiar and Waldfogel 2018b). For example, Zhang (2018) shows that making licensed content more usable (by removing digital rights management restrictions) increases sales of that content. Studies show that streaming services are associated with lower sales in conventional channels (Hiller 2016, Wlömert and Papies 2016, Aguiar and Waldfogel 2018b, Datta et al. 2018) but lead to higher per-capita consumption (Datta et al. 2018), so the aggregate revenues of the music industry remain more or less unchanged (Wlömert and Papies 2016, Aguiar and Waldfogel 2018b). However, some studies also suggest complementary effects of streaming. For example, Nguyen et al. (2014) find that streaming services do not affect physical purchases but increase attendance at live concerts. Online video in particular has been shown to indirectly increase sales of complementary products in the context of fashion retail (Kumar and Tan 2015). For music, however, there may be a more direct effect on sales. In October 2008, YouTube rolled out click-to-buy links to iTunes and Amazon Next to music videos. Rights holders can choose to add such links to their own uploads and to user-generated videos for which the rights holder is identified through YouTube's Content ID system (see YouTube 2008). Given the mixed prior evidence and the specific institutional details, however, the average effect of YouTube on sales of individual songs remains an empirical question.[4]

2.2. User-Generated Content

YouTube is an open licensed platform, meaning that anyone can provide content to the platform, and the platform has taken steps for rights holders to be compensated for the use of their content. Therefore, YouTube content linked to a particular song is usually a combination of official videos and UGC, which may affect its impact on recorded music sales.

YouTube is an immensely popular music platform. More than 80% of YouTube visitors use it for music (International Federation of the Phonographic Industry 2016), and video-based streaming accounted for more than 50% of the 317.3 billion music streams in the United States in 2015.[5] Approximately 30% of YouTube videos (and 40% of total views) are music videos (Liikkanen and Salovaara 2015). YouTube is not heavily invested in the creation of its own music content but provides financial incentives to third parties to upload content (Tang et al. 2012). Consistent with evidence that ad revenue sharing and content commercialization shift incentives toward creating more mass-oriented content (Sun and Zhu 2013), the introduction of the so-called partner program in 2007 quickly led major record labels to make their music video libraries available on YouTube (see TechCrunch 2007). However, most content on YouTube is user generated. Consumers upload videos of live performances, videos that embed the lyrics of a song, or derivative works such as cover versions and parodies (Liikkanen and Salovaara 2015). Through YouTube's Content ID technology, all videos are matched to a database of copyrighted material. When uploads are classified as infringing, rights holders can choose between blocking the infringing video and "monetizing" the content by sharing revenue from advertisements. In testing, YouTube reported that 90% of claims created through Content ID led to rights holders choosing monetization (see King 2008). Although this gives substantial control to rights holders compared with closed platforms, it is more difficult for artists or labels to entirely withdraw content from YouTube because new UGC is constantly being uploaded.[6]

Overall, little is known about the relationship between original innovation and adaptations and derivative works. The literature has focused on related yet different questions. A stream of work studies how user-generated knowledge affects firm innovation (for a recent example, see Gambardella et al. 2017). Empirical evidence on the role of intellectual property (IP) rights shows that IP protection (or stronger enforcement) hinders follow-on innovation, measured as subsequent academic knowledge creation (Williams 2013), developer activity in open-source software projects (Wen et al. 2013), and the number of patent citations (Galasso and Schankerman 2015). Hence, UGC (especially adaptations and derivative work) may act complementarily to sales of official content, or consumers may perceive such content as a substitute. Regarding online music videos and record sales, this implies that the fact that YouTube is an open platform can shape its impact on recorded music sales, and it is not clear in which direction.

2.3. Digital Distribution Platforms and Effects on Consumption Patterns

Internet-enabled distribution platforms often offer a variety of tools to aid content discovery. The reduction in search cost has the potential to affect the sales distribution (Anderson 2006). A first example is aggregate top lists of the best-selling products in a given product category, week, and geographic area. Aggregate popularity information (observational learning) can drive concentration (Salganik et al. 2006, Sorensen 2007, Cai et al. 2009, Hinz et al. 2011) but also can benefit niche content when the same level of popularity implies higher quality of niche versus broad-appeal products (Tucker and Zhang 2011).

More personalized recommendations can come directly from peers (often encouraged by platforms making it easy to share content) or from algorithms using social data. For example, evidence shows that social interactions determine which videos on YouTube become successful (Susarla et al. 2012). Theoretical work argues that consumers with less common preferences may benefit more from social recommendations, which trigger demand for niche products (Hervas-Drane 2015). This is supported by evidence that the popularity distribution of e-commerce sales can shift toward the tail (Oestreicher-Singer and Sundararajan 2012a). Although search and recommendation tools guide consumers to content they were previously unaware of, they can also increase the number of consumed units and the chance of consuming content that many others are also consuming (Hosanagar et al. 2014). In our empirical context, it is not clear how this trade-off plays out. Because musical content is heterogeneous in its breadth of appeal and prior exposure, we also investigate whether the effects differ when we distinguish between niche content (realized appeal) and new content (ex ante unknown appeal).

3. Institutional Background

3.1. Variation from YouTube and the GEMA Shock

YouTube is a unique setting to study the relationship between the availability of online music videos and recorded music sales. Although YouTube has contracts with rights holders in most countries, the question of compensation was subject to a long-standing legal dispute between YouTube and GEMA in Germany. GEMA is the state-authorized (de facto monopolist) collecting society and performance rights organization in Germany.[7] Collecting societies are bodies that ensure that royalties from any kind of reproduction (e.g., physical and digital reproduction, public performance, etc.) reach artists and publishers, making them important institutions for artists because royalties are a major part of income, independent of any private contracts with record labels (Kretschmer 2005). A large international network of sister collection societies represents the rights of German artists/publishers in international markets, and GEMA does the same for international artists/publishers in the German market. That is, virtually every professional musician is either directly or indirectly a member of GEMA, which led to the so-called GEMA presumption, a case law presuming that rights of all musical works are managed by GEMA in Germany.[8]

The expiry of an initial agreement between YouTube and GEMA in 2009 triggered renewed negotiations about the appropriate level of compensation. In fear of high subsequent (and retrospective) payments, YouTube began blocking music videos on April 1, 2009 (see O'Brien 2009). The left-hand panel of Figure 1, showing Google Trends search volume for the term "gema" from April 2008 to April 2010, indicates a spike in the week when the blocking began but not much systematic movement before and after. This suggests that the shock came unexpectedly to consumers and most artists, publishers, and record labels.[9] The situation persisted until November 1, 2016, when YouTube and GEMA finally announced that they had reached an agreement and the restrictions on music videos were lifted.[10]

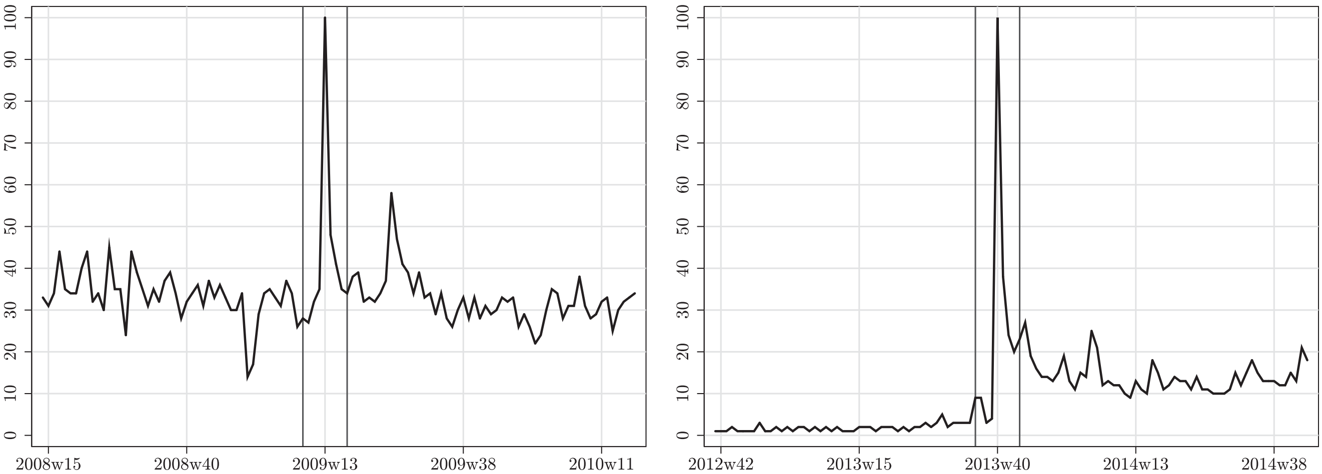
However, this does not necessarily imply that German YouTube users did not have access to any music videos. Publishers/artists can negotiate independent contracts with any online or offline licensee, so publishers and artists may drop out of GEMA to reach individual agreements with YouTube in Germany.[11] However, this may not be optimal. First, royalty income from the collecting society comes from a variety of digital (e.g., download and streaming services) and physical (e.g., radio broadcasting and public performance) sources. According to GEMA's annual report, online rights accounted for less than 2% of the overall combined income from online, physical duplication, and radio and television in 2008.[12] Hence, income from digital distribution may be too small an amount to forgo all other royalty income for most artists. Second, by joining a collecting society, individuals benefit from reduced contracting costs and increased bargaining power. This is even more beneficial for members of international collecting societies, where it can be especially costly to negotiate with various potential licensees abroad across different legal systems. To avoid this potential endogeneity of selecting into (or out of) YouTube in our estimates, we focus on a very short time window of four weeks before and after the blocking began on April 1, 2009.[13]

3.2. Long-Run Supply-Side Reactions and the Launch of Vevo

The royalty dispute triggered a controversial discussion in the German music industry. Whereas some artists agree with the position of GEMA because they believe that YouTube's royalty rates are too low, others simply want their videos to be seen.[14] Representatives of Sony Music and Universal Music have publicly criticized GEMA for not working harder toward an agreement (see Spahr 2011).

Record labels are not members of GEMA (they do not create music) and therefore do not receive any royalty income. However, on top of a potential positive effect on record sales, they directly benefit from advertising revenues generated by YouTube. Not surprisingly, therefore, record labels are heavily

Figure 1. Google Search Volume for GEMA and Vevo



Source. Google Trends.
Notes. The figure shows relative Google search volumes for "gema" and "vevo" in Germany for April 2008–April 2010 (left panel) and October 2012–October 2014 (right panel), respectively. Vertical lines indicate the sample period for the econometric analysis.

invested in online music video. Sony Music and Universal Music, with a joint market share of more than 46% in 2012, for example, hold majority stakes in the music video service Vevo.[15] Since its launch in 2009, Vevo has been partnering with YouTube in most countries. Accordingly, 97% of its 51.6 million unique viewers accessed Vevo content through YouTube in December 2012, making Vevo the most viewed channel on YouTube, accounting for a third of all unique viewers on YouTube (see Comscore 2013). As a workaround for the GEMA–YouTube deadlock, Vevo negotiated its own licensing deal with GEMA and launched the dedicated platform Vevo.com—with content hosted outside YouTube—in the German market on October 1, 2013 (see Cookson 2013). Overnight, 75,000 music videos became available on the German internet. We use this event as an auxiliary test for our general findings at a different time (2009 versus 2013), with a different scope (all musical content versus content by two major labels), in a different direction (making videos available versus unavailable), and on a different type of platform (an open versus closed platform that does not allow for UGC).

4. Data

We construct a unique data set by matching sales information to song- and artist-specific metadata and measures of online video availability from a variety of sources. Table 1 gives a summary and definitions of all variables used in this paper; descriptive statistics of the key variables are in Table 2.

4.1. Sales Data

Sales data come from GfK Entertainment, a market research firm collecting weekly (scanner) data. The data include information from 50 (online and offline) retail outlet chains and 27 digital retailers in Germany and Austria that collectively represent more than 90% of all retailers.

We have access to the weekly number of units sold for the subsample of all songs that were among the 1,000 highest grossing songs (based on cumulative sales across all distribution channels) at least once in weeks 10–23 of 2009 in the physical channel (in Germany and Austria) and the digital download channel (in Germany). We also have sales information for weeks 36–44 of 2013. For 2013, we observe physical sales (in Germany and Austria) and digital downloads (in Germany and Austria), and we can distinguish between the number of weekly streams via free services and subscription-based services (in Germany). Note that YouTube and Vevo plays are not included in the latter variable. For most analyses, we use a four-week window around the respective experiments.

The data cover sales of individual songs. In the physical sales channel, this is equivalent to the sales of singles, which implies that the song has been released on a physical medium—and stocked in brick-and-mortar outlets—for us to observe positive sales figures.[16] In the digital channel, our data capture sales of songs independent of whether they are released as stand-alone products (digital single) or as part of an album. This is because consumers can (almost) always buy any individual song on an album in digital record stores. This is reflected in the share of digital song sales in our data; on average, it is 79% in 2009 and 97% in 2013. Looking at aggregate sales figures covering the top 1,000 best-selling albums in the German market in the same period, the digital share of top 1,000 album sales is 5% in 2009 and 18% in 2013. This is in line with industry reports in which the digital share in overall sales figures (songs and albums, entire

Table 1. Variable Definitions

Variable	Definition
Dependent variables	
Total (ln + 1)	Weekly number of sold units (physical plus digital)
Physical (ln + 1)	Weekly number of physically sold units
Digital (ln + 1)[a]	Weekly number of digitally sold units
Streaming: Total (ln + 1)[b]	Weekly number of streams (free plus premium)
Streaming: Free (ln + 1)[b]	Weekly number of streams on free streaming services (not including YouTube and Vevo)
Streaming: Premium (ln + 1)[b]	Weekly number of streams on subscription-based streaming services
Independent variables	
After (0/1)	Weeks after 2009 week 14/2013 week 40 (GEMA shock/entry of Vevo)
Video (0/1)	(At least one) song-specific video on U.S. YouTube uploaded prior to April 1, 2009, by Vevo
Germany (0/1)	Data from Germany
Vevo Label (0/1)	Universal Music or Sony Music (and their subsidiaries)
UGC: Small Share (0/1)	Share of official song-specific videos is bigger than or equal to that of the average song in the sample
UGC: Official and UGC (0/1)	Share of official song-specific videos is smaller than that of the average song in the sample
UGC: Only UGC (0/1)	Only user-generated song-specific videos
Newcomer: Two Months (0/1)	Earliest release of an artist not more than two months before GEMA shock
Newcomer: No Album (0/1)	Artist has never released an album before the GEMA shock
Newcomer: First Year (0/1)	Earliest release of an artist not more than one year before GEMA shock
Niche: Never U.S. (0/1)	Artist never appeared on the U.S. charts (album and single), 2000 week 1–2009 week 14/2000 week 1–2013 week 36
Niche: German (0/1)	Artist has German origin
Niche: Genre (0/1)	Song is not in a mainstream genre (i.e., not pop or rock)

[a]Data are not available for Austria in 2009.
[b]Data are available only for Germany in 2013.

German market) is 11% in 2009 and 23% in 2013.[17] We prefer to run our analysis at the song level because music videos are directly linked to individual songs, not entire albums. In fact, as we show in Section 6.3.1, the choice of level of aggregation is an important aspect when comparing our results with those of the prior literature.

Comparing the head and tail of the top 1,000 list shows that it covers a very large fraction of the market. In our data, songs at the bottom of the list never sell more than 12 units (physical and digital combined) in a given week in the 2009 sample and never more than 20 units/12 streams in the 2013 sample, whereas the top song sells, on average, 37,871 units in 2009 and 28,121 units and 543,653 streams in 2013. Hence, demand for songs that do not make it to the top 1,000 list is small in both relative and absolute terms. However, we do not only track songs while they are part of the top 1,000 list. For each song that appears in the top 1,000 list at some point in our observation period, our data also let us observe sales figures in weeks where the song is not included in the top 1,000 list but available for purchase or streaming.

4.2. Metainformation

We match metainformation such as the release date and information on other releases of the same artist, genre, geographic origin of the artist, and record label using data from MusicBrainz, an online platform for music enthusiasts. MusicBrainz contains user-generated information on approximately 20.5 million songs. Crucial information for our analysis, especially exact release dates (not only month or year), are not available for every song, but in some cases, we found additional information from other sources, such as iTunes, Wikipedia, and Discogs. Data on prior international success of an artist come from historical chart rankings in the United States (week 1 of 2000–week 36 of 2013) from Billboard.

We note some technical details. We do not observe unique common identifiers across the various databases, which is why we rely on a comparison of text strings, that is, artist names and song titles. Because the additional data sets are very large and we need to match multiple ones, going through each combination of potential matches is close to impossible. Furthermore, variations in artist names and song titles (e.g., "featuring," "feat.," "(Radio Version)," "(Club Mix)") make one-to-one matching too restrictive, potentially causing (too many) matching errors. Because manual inspection of match candidates is potentially error prone and there is no structured way to quantify error rates, we develop a statistical matching algorithm.[18] We compare pairs of potential matches along a number of metrics (such as the Levenshtein distance or Soundex), manually code a random subsample (training data, n = 7,654) to estimate parameters of a logit model, and use those parameters to estimate match probabilities in the full sample. From the statistical properties of the underlying model and training data, we can estimate error rates. Our model performs at an estimated rate of 4.4% of type I errors and 4.6% type II errors. This is much more precise

Table 2. Descriptive Statistics

	No video		Video		Total	
Variable	Mean	SD	Mean	SD	Mean	SD
Sample of the GEMA shock, 2009 weeks 10–18						
log(*Total* + 1)	1.783	2.277	3.575	2.168	2.764	2.390
log(*Physical* + 1)	0.507	1.306	0.910	1.675	0.728	1.532
log(*Download* + 1)	1.673	2.229	3.482	2.158	2.664	2.369
After	0.451	0.498	0.449	0.497	0.450	0.497
Video	0	0	1	0	0.548	0.498
UGC: Small Share	0	0	0.278	0.448	0.152	0.359
UGC: Official and UGC	0	0	0.221	0.415	0.121	0.326
UGC: Only UGC	0	0	0.501	0.500	0.275	0.446
Newcomer: Two Months	0.028	0.165	0.045	0.206	0.037	0.189
Newcomer: No Album	0.046	0.210	0.040	0.195	0.043	0.202
Newcomer: First Year	0.065	0.247	0.065	0.246	0.065	0.247
Niche: Never U.S.	0.821	0.383	0.642	0.480	0.723	0.448
Niche: German	0.616	0.486	0.470	0.499	0.536	0.499
Niche: Genre	0.469	0.499	0.355	0.479	0.407	0.491
Number of songs	700		842		1,542	
Number of artists	535		549		999	
Observations	6,202		7,509		13,711	
Sample of the Vevo entry, 2013 weeks 36–44						
log(*Total* + 1)	3.333	2.058	3.869	2.030	3.576	2.062
log(*Physical* + 1)	0.199	0.796	0.207	0.835	0.203	0.814
log(*Download* + 1)	3.315	2.065	3.857	2.031	3.561	2.068
log(*Total Streams* + 1)	4.776	4.079	5.583	3.908	5.142	4.023
log(*Free Streams* + 1)	3.150	4.118	3.726	4.240	3.411	4.184
log(*Premium Streams* + 1)	4.547	3.832	5.323	3.669	4.899	3.779
After	0.588	0.492	0.574	0.495	0.582	0.493
Video	0	0	1	0	0.454	0.498
Vevo Label	0.016	0.127	0.021	0.144	0.019	0.135
Number of songs	937		759		1,696	
Number of artists	631		485		1,027	
Observations	7,962		6,614		14,576	

Note. SD, standard deviation.

than the heuristic approaches that are commonly used in the related literature (e.g., a reduction of 16 percentage points in the type I error rate in the best-performing method discussed in Raffo and Lhuillery (2009, figure 4)). Our final data set includes 1,542 songs from 999 artists.

4.3. Music Video Data

To build a song-level measure of video availability on YouTube, we would ideally observe which songs had corresponding videos on the German YouTube just before the ban on April 1, 2009. Because such historical data are not available, we construct a proxy by gathering the first 20 search results (this reflects the first page of search results on YouTube) from a query of artist name and song title on the U.S. version of YouTube (using YouTube's application programming interface).[19] In many cases, not all videos that YouTube returns for a song are directly related to that song. Sometimes we observe videos related to other songs of the same artist, songs from similar artists, etc. We treat videos as relevant if the video title includes the artist's name and at least three words of the song title. Using the upload date of each thus defined video, we construct our measure of availability. We set the dummy variable $Video_i$ to one if at least one video corresponding to song i was uploaded before April 1, 2009. Identification of the YouTube effect will thus come from differences between songs that had corresponding videos on YouTube and those that did not.

We use data from the U.S. YouTube because we can realistically assume that the German YouTube would offer the same content as the U.S. YouTube had the GEMA shock not happened. A simple plausibility check for the latter is to compare U.S. search results with those from Austria, Germany's neighbor, which shares the same language and similar culture but was not affected by the GEMA shock. George and Peukert (2014) conduct such an exercise with a random selection of almost 1,000 songs released between 2006 and 2011, collecting search results on the German,

Table 3. YouTube in the United States, Germany, and Austria

Country	Share of directly relevant videos	Share of total views	Official video share
United States			
Mean	0.7755	0.8250	0.0868
Standard error	0.0014	0.0014	0.0025
Austria			
Mean	0.7726	0.8213	0.0893
Standard error	0.0014	0.0014	0.0026
Germany			
Mean	0.7485	0.7483	0.0502
Standard error	0.0016	0.0021	0.0020

Source. George and Peukert (2014).
Notes. Results are based on the top 20 YouTube search results for 950 randomly selected songs released between 2006 and 2011 (based on data from MusicBrainz). The search was carried out on August 21, 2014. Relevancy is defined as a YouTube video title containing the artist name and at least three words of the song title. Total views are calculated as the cumulative number of views of all 500 videos shown on the first 20 result pages. Official videos are identified by the word "official" in the title or uploader name.

Austrian, and U.S. versions of YouTube for each song. Table 3 shows that the share of directly relevant videos on the first result page (defined as earlier) is not significantly different in the United States and Austria but 2.7 percentage points lower in Germany. The share of total views (sum of views of all 500 videos on the first 20 result pages) on the first result page in the United States and Austria is about 82% (no significant difference). This share is about 75% in Germany. Almost 9% of the relevant videos on the first result page are official videos in the United States and Austria (no significant difference); in Germany, this share is only 5%.[20] This suggests that the Austrian version of YouTube looks very much like the U.S. version of YouTube, whereas top search results on the German YouTube are clearly different.

In essence, therefore, we have a measure of video availability just before the ban, based on the U.S. YouTube, not the German YouTube, which, however, is likely to be essentially the same before April 2009. For 54% of the songs in our sample, there is at least one corresponding YouTube video that predates the GEMA shock. The seemingly high share of songs without videos is plausible because it is likely that a substantial number of artists and record labels believed that YouTube displaces record sales and did therefore not upload official content and actively sent takedown requests for UGC. In the online appendix, we report lists of the top 20 (based on their peak rankings in our observation period) songs with videos and without videos. Songs without videos tend to have lower peak ranks than songs with videos. Consistent with this, descriptive statistics in Table 2 show that songs with videos have substantially higher average sales.[21] However, as long as this difference is constant, it will cancel out in the difference-in-differences model. In the econometric model, we control for unobserved factors that may be correlated with such differences via song fixed effects and week fixed effects and show that the identifying assumption of the difference-in-differences model holds.

Similarly, we would like to observe which songs had corresponding videos on the German Vevo website when it was launched on October 1, 2013. Vevo does not provide such a list, but we can make use of the fact that Vevo is part of YouTube in many other countries, including the United States. The underlying assumption is that the content on Vevo's standalone German platform is the same as Vevo's content on YouTube.[22] We take advantage of the fact that Vevo uses artist-specific usernames to upload videos to YouTube. For example, the corresponding username for official videos by Justin Bieber is JustinBieberVevo. Accordingly, we define a song as having a Vevo video if at least one song-specific video is uploaded by Vevo, which is the case for 37% of the songs in our sample.[23]

To identify official non-Vevo videos, we manually went through the official YouTube accounts of artists and record labels and flagged videos as official if they were uploaded by these accounts. The average share of official videos among the relevant YouTube results is 6%. Accordingly, the average share of user-generated videos is 94%. Casual inspection of those videos suggests that these are mostly live and lyric videos, remixes and cover versions, and a very few parodies (see Liikkanen and Salovaara 2015).

4.4. Heterogeneity

We distinguish between established and newcomer artists in several ways using information on all historical releases of an artist. First, we define *Newcomer: Two Months* as a dummy variable indicating whether the first release of an artist did not appear before February 1, 2009. We observe 59 songs in this category. Similarly, *Newcomer: First Year* is defined as an indicator of whether the first release of an artist did not appear before April 1, 2008. This identifies 102 newcomer songs. Our third variable, *Newcomer: No Album*, indicates whether the artist never released an album before the GEMA shock, that is, April 1, 2009. The number of newcomer songs for this definition is 68. Prominent examples of newcomers according to these definitions are Oceana, Steve Appleton, Katy Perry, and Lady Gaga.

We then define three empirical measures of mainstream versus niche music. First, we use historical sales data from the United States (*Billboard* Top 100 single and album charts) to define a measure of international success. Examples of German artists that appeared at least once on the *Billboard* charts include Cascada, Sarah Connor, and Tokio Hotel. However, consistent with the evidence in Ferreira and Waldfogel (2013) showing that Germany almost exclusively exports music to Austria and Switzerland, these are exceptions. Accordingly, our first measure of niche artists consists of 638 artists who never appear in the U.S. top charts, which we include in the category *Niche: Never U.S*. Second, based on data about their geographic origin, we classify 601 artists as *Niche: German*. This represents 60% of all artists in our data, which again is consistent with the estimate of the domestic share of music consumption in Ferreira and Waldfogel (2013). Third, we identify the best-selling genres by looking at cumulative sales in our data. Pop and rock account for 74% of overall sales in the German market before the GEMA shock such that we include all other genres in the category *Niche: Genre*.

One might be concerned that songs of newcomer or niche artists differ in their video availability. The descriptive statistics in Table 2 do not confirm this. Applying Bayes' theorem, we see that the probability that a song does not have a video if it is from a niche artist is roughly 50%, independent of how we measure *Niche*.[24] The probability that a song does not have a video if it is from a newcomer artist is between 34% and 48%, depending on the definition of *Newcomer*.

5. Empirical Specification and Results

We first introduce and discuss the identification strategy and then report our baseline results on the average effect of online music videos on recorded music sales. We go on to test whether this effect is driven by differences in the song-specific amount of official content and UGC on YouTube. Finally, we investigate heterogeneity and distinguish between new artists and niche artists. We report results of additional analyses (some of which are described in more detail in the online appendix) that largely support the robustness of our results and help us rule out alternative explanations.

5.1. Identification Strategy

We identify the effect of music videos on sales of recorded music using exogenous variation from removing access to videos on YouTube in a difference-in-differences model. Essentially, we compare sales of songs *with* videos with sales of songs *without* videos before and after the natural experiment. Our baseline specification can be written as

$$\log(Sales_{it}^{k} + 1) = \alpha + \sum_{t} \beta^{t} w_t + \delta(After_t \times Video_i) + \mu_i + \varepsilon_{it}, \qquad (1)$$

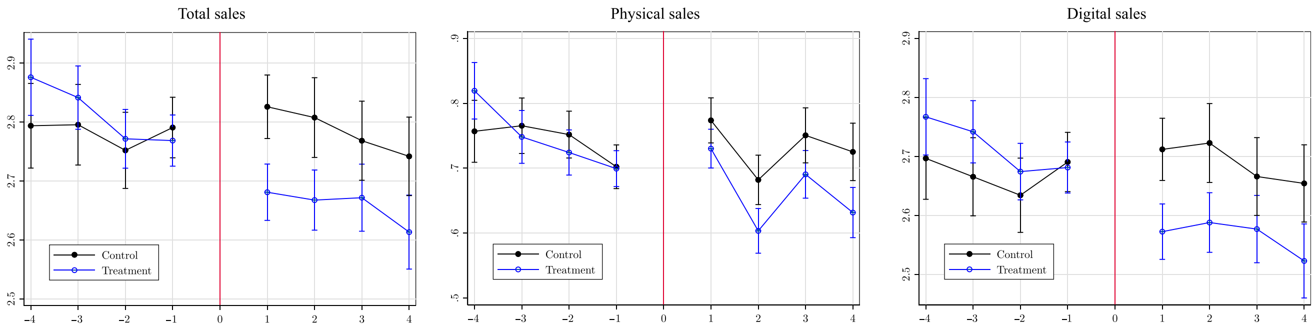
where $Sales_{i,t}^{k}$ are unit sales of song i via channel k (physical, digital) in week t, $Video_i$ is our measure of video availability and defines the treatment group, and $After_t$ indicates the time period after the GEMA shock. Our estimate of the causal effect of the experiment is δ. We further include week fixed effects β^t and song fixed effects μ_i to control for unobserved song-specific, time-invariant and time-specific, song-invariant heterogeneity. Because of these fixed effects, we implicitly control for but cannot separately identify coefficients for $Video_i$ and $After_t$. Our preferred specification reports standard error estimates clustered at the song level.

The identifying assumption in any difference-in-differences setting is that the treatment and control groups would have followed similar trends in the dependent variable had the policy shock not happened. A necessary condition for this assumption to hold is that trends in the dependent variables of the treatment and control groups are parallel before the experiment. To see whether this condition holds, we can plot a sales measure over time. In Figure 2, we partial out song fixed effects and plot averaged residuals for each group and each week regarding total sales, physical sales, and digital sales. The plot shows that the treatment and control groups follow similar trends before the GEMA shock and start to diverge substantially afterward. Noting that the 90% confidence bands (calculated using the standard error of the mean) overlap in the preshock period, we can conclude that the parallel trends assumption is consistent with our data. More formally, we can test whether the difference in the dependent variable across points in time is zero in the preshock period, as is often done in the literature (for one of the first applications of this idea, see Autor 2003). We estimate a model defined as

$$\log(Sales_{it}^{k} + 1) = \alpha + \sum_{t} \beta_0^{t} w_t + \sum_{t} \beta_1^{t} (w_t \times Video_i) + \mu_i + \varepsilon_{it}, \qquad (2)$$

in which we can test, week by week, whether treatment and control groups differ in their sales dynamics

Figure 2. (Color online) Trends of Treatment and Control Groups Before and After the GEMA Shock



Notes. The vertical axis in each panel shows the average demeaned total/digital/physical sales, that is, averaged residuals $\overline{\hat{y}_{vt}} = \overline{\hat{y}_{it} - \hat{\mu}_i}$ derived from the model $y_{it} = \log(Sales_{it} + 1) = \alpha + \sum_t \beta_0^t w_t + \sum_t \beta_1^t (w_t \times Video_i) + \mu_i + \varepsilon_{it}$ for $Video_i = 0$ and $Video_i = 1$. The horizontal axis shows weeks before and after April 1, 2009. Lines with filled markers (control group) show the average sales of songs without at least one video uploaded to the U.S. YouTube before April 1, 2009. Lines with hollow markets (treatment group) show the average sales of songs with at least one video uploaded to the U.S. YouTube before April 1, 2009. Error bars indicate the 90% confidence bands (standard error of the mean).

($H_0 : \beta_1^t = 0$). Table 4 shows estimates of β_1^t coefficients for total sales, physical sales, and digital sales in the preshock period. Across almost all columns and preexperiment weeks, we cannot reject the hypothesis that the difference in sales in the treatment and control groups is equal to zero. Although these results are reassuring, we discuss an alternative identification strategy (using cross-country variation) and a falsification exercise (placebo country and timing) after the baseline results below.

5.2. Displacement or Promotion?

5.2.1. Baseline Results. Our baseline results are reported in Table 5.[25] In columns (1) and (5), we report the results of an aggregated model specification, where we look at total sales of a song as an average in the pre- and postshock periods. This specification lets us address the potential issue that serial correlation may lead to incorrect inference (Bertrand et al. 2004). We further report results of different methods of estimating standard errors (clustered at the artist level in column (1) and clustered at the song level in column (5)). The estimated difference-in-differences coefficient is negative and significant at the 5% level. The point estimate is −0.142, which implies a decrease of 13%.[26]

Turning to the preferred disaggregated sample in columns (2) and (6), we get very similar results. The point estimate translates into a percentage reduction of 14% of total sales, with the 90% confidence interval (CI) between −21% and −6%. Because it is likely that the temporal correlation structure of sales of the same song is stronger than those of sales of different songs by the same artist, we continue to report results with standard errors clustered at the song level in the rest of this paper.

In columns (3) and (4) and columns (7) and (8), respectively, of Table 5, we distinguish between physical and digital sales of the same song. Although the point estimate for physical sales is half the size of the point estimate for digital sales (−7% versus −14%), the coefficients are not statistically different from each other in the sense that 90% CI bands overlap. Hence, we conclude that YouTube's average effect on sales of songs is not statistically different regarding physical and digital sales.

5.2.2. Alternative Identification Strategy and Falsification Exercises. An alternative identification strategy that does not use song-specific information on video availability is to compare sales in Germany with sales in a different country not affected by the GEMA shock. Austria, Germany's neighbor, which shares the same language and similar culture, is a prime candidate. In this exercise, we assume that the GEMA

Table 4. Group Differences in the Period Before the GEMA Shock

	(1)	(2)	(3)
Variable	Total	Physical	Digital
$t_{-4} \times Video$	0.113*	0.075	0.097
	(0.068)	(0.046)	(0.066)
$t_{-3} \times Video$	0.091	−0.004	0.117*
	(0.062)	(0.042)	(0.060)
$t_{-2} \times Video$	0.036	−0.018	0.055
	(0.058)	(0.036)	(0.056)
$t_{-1} \times Video$	−0.013	−0.003	0.000
	(0.046)	(0.029)	(0.045)
Observations	7,543	7,543	7,543
$\overline{R^2}$	0.922	0.920	0.924

Notes. The dependent variable is (log + 1) *Weekly Sales* in units. The term *Video* indicates that at least one song-specific video was uploaded on the U.S. YouTube prior to April 1, 2009. Sample only includes weeks before the GEMA shock. Song and week fixed effects and the constant are not reported. Standard errors are in parentheses, clustered at the song level.

*$p < 0.10$.

Table 5. Baseline Results: Average Effect of the GEMA Shock on Song Sales

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variable	Total (avg)	Total	Physical	Digital	Total (avg)	Total	Physical	Digital
After × *Video*	−0.142**	−0.146***	−0.072*	−0.153***	−0.142**	−0.146***	−0.072*	−0.153***
	(0.068)	(0.050)	(0.037)	(0.049)	(0.070)	(0.052)	(0.037)	(0.051)
Observations	3,084	13,711	13,711	13,711	3,084	13,711	13,711	13,711
SE cluster	Song	Song	Song	Song	Artist	Artist	Artist	Artist
$\overline{R}^2$	0.918	0.892	0.877	0.894	0.918	0.892	0.877	0.894

Notes. The dependent variable in columns (1) and (5) is the average log(1 + *Weekly Total Sales*) of song *i* in the preshock period and postshock period. For all other columns, the dependent variable is log(1 + *Weekly Total*/*Physical*/*Digital Sales*) of song *i* in week *t*. The variable *Video* indicates that at least one song-specific video was uploaded on the U.S. YouTube prior to April 1, 2009. The variable *After* indicates weeks after April 1, 2009. All models include song fixed effects, and in columns (2)–(4) and (6)–(8), we additionally include week fixed effects. The constant is not reported. Standard errors (SEs) are in parentheses, clustered at the song/artist level. avg, Average.
$^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

shock affected all songs in the same way in the German market, that is, that all songs have corresponding videos on YouTube that are all blocked. In Table 6, we report results of a model that compares sales of a song in Austria with its sales in Germany, before and after the GEMA shock. Note that digital sales data do not exist for Austria, so we can compare only physical sales. The number of observations is accordingly twice as large as in the disaggregated specifications (columns (2)–(4) and (6)–(8) of Table 5). Because we have two observations per song, we can identify a country coefficient. Given that Germany is a larger market than Austria, this coefficient is positive and significant, as expected. The difference-in-differences coefficient *After* × *Germany* is negative and significant. The effect size is −7% (90% CI, [−12%, −1%]).

We also perform two falsification exercises. First, we estimate placebo versions of our model pretending that the GEMA shock took place either two weeks before or two weeks after it actually did. As illustrated in Figure 3, we split the sample of our baseline model (running from t_{-4} to t_{+4}) in two parts and estimate models on the two subsamples running from t_{-4} to t and from t to t_{+4}, setting the respective dates of the placebo experiments to t_{-2} and t_{+2}; that is, we run two types of placebos. For the first, if there is no general underlying trend, we should expect an effect that is close to zero. For the second, we expect an effect that is close to zero only if we assume that the effect of the real experiment is instantaneous and constant. Hence, we can take this exercise as a test of whether the effect of the true experiment is immediate and whether it changes in the observed time frame.[27]

Table 6. Alternative Identification Strategy

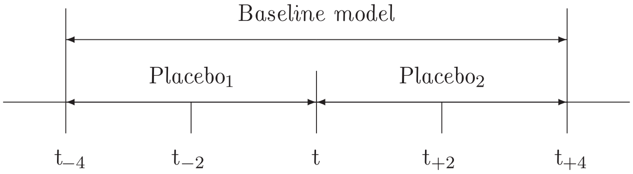
Variable	(1)
Germany	0.424***
	(0.046)
After × *Germany*	−0.068**
	(0.033)
Observations	27,422
$\overline{R}^2$	0.462

Notes. The dependent variable is (log + 1) *weekly physical sales* in units (digital sales are not available for Austria). Song and week fixed effects and the constant are not reported. Standard errors in are parentheses, clustered at the song level.
$^{**}p < 0.05$; $^{***}p < 0.01$.

In the results reported in columns (1) and (2) of Table 7,[28] we define the after period to include the week of the placebo experiment, but we get very similar estimates if we treat the placebo experiment week as part of the before period. The coefficients of *After* × *Video* in columns (1) and (2) of Table 7 are small and not significantly different from zero, suggesting that our results are not driven by a general trend that started before the GEMA shock and that the effect of the shock persists. Table 7 reports only results concerning total sales as the dependent variable, but we also do not find significant effects when we distinguish between physical and digital sales. Second, we run a country-based placebo exercise and estimate our model on data from Austria. If our results are driven by confounding temporal variation that coincides with the GEMA shock, we should see a similar effect in Austria. Alternatively, because music videos on YouTube remained available for Austrian consumers, we should not see any change in the Austrian sales of songs that are affected by the experiments in Germany. This is indeed what the data tell in column (3) of Table 7. We find that songs with videos on YouTube do not have significantly different sales in Austria compared with songs without videos on YouTube before and after the GEMA shock in Germany. The coefficient is very close to zero.

Although our results are robust to a number of specifications, some concerns regarding data structure

Figure 3. Falsification Exercises: Timing of Placebo Experiments



and measurement error may remain. In the online appendix, we speculate on the possible effects of measurement error, show that our results are broadly robust to different estimation windows, and conclude that the results are not likely to be driven by price changes.

5.3. The Role of Content Type and Platform Type

One of the specific features of YouTube as a licensed open platform is that it can host both official content and UGC. We provide two tests of whether this distinction can lead to different effects on sales. First, as described earlier, we separate the available song-specific videos by official content and UGC. Second, we make use of the fact that Vevo is a closed platform, where only rights holders themselves can contribute, and estimate how its launch affected sales of recorded music across different distribution channels.

5.3.1. User-Generated Content. Table 8 reports results from a specification where we define dummy variables based on the share of official content and UGC. We categorize songs based on the amount of official content and UGC on YouTube. On average, the share of official videos in the directly related search results is 6%. The median is zero and the 75th percentile is 10%. We define a song as having a *Small UGC Share* if the share of official videos is larger than 10%, which is roughly equivalent to two official videos on YouTube's first result page.[29] Songs without any official videos are categorized as *Only UGC*. Accordingly, songs with a share of official videos of between 0% and 10% are classified as *Official and UGC*. Similar to the baseline specification earlier, the omitted category is comprised of songs without videos. We report results for total sales (column (1)), physical sales (column (2)), and digital sales (column (3)). The point estimates of *After* × *Small UGC Share*, *After* × *Official and UGC*, and *After* × *Only UGC* are very similar. We interpret these results as evidence that the effect of music videos on record music sales operates in a similar fashion independent of whether there is more or less UGC on YouTube. In the online appendix, we show that the parallel trends assumption is supported by the data. The pre-GEMA shock sales dynamics of treatment (*Small UGC Share*, *Official and UGC*, *Only UGC*) and control songs do not systematically vary by the amount of user-generated video content on YouTube.

Table 7. Falsification Exercises: Placebo Experiments

	Timing		Country
	−2 weeks	+2 weeks	Austria
Variable	(1)	(2)	(3)
After × *Video*	−0.034	−0.037	0.001
	(0.060)	(0.043)	(0.001)
Observations	4,144	3,852	13,711
$\overline{R}^2$	0.803	0.734	0.755

Notes. The dependent variable is (log + 1) *Weekly Total Sales* in units, in Germany (columns (1) and (2)) and Austria (column (3)). Column (3) includes only physical sales because digital sales data are not available for Austria. The term *Video* indicates that at least one song-specific video was uploaded on the U.S. YouTube prior to April 1, 2009. Song and week fixed effects and constant are not reported. Standard errors are in parentheses, clustered at the song level.

5.3.2. The Entry of Vevo. The GEMA shock was partly reversed when Universal Music and Sony Music launched their own music video platform, Vevo, on October 1, 2013. Although it is independently interesting to ask whether the availability and unavailability of music videos affect sales to comparable magnitudes (but with opposite signs), the entry of Vevo lets us speak to the role of UGC from another angle. Vevo only hosts official videos, whereas YouTube hosts a large share of UGC videos. Table 9 replicates our baseline regressions with data covering four weeks before and four weeks after the Vevo launch. As described in Section 4.3, we can measure whether song-specific music videos were available on

Table 8. Official vs. User-Generated Content

	(1)	(2)	(3)
Variable	Total	Physical	Download
After × *Small UGC Share*	−0.183***	−0.085	−0.199***
	(0.069)	(0.062)	(0.066)
After × *Official and UGC*	−0.146*	−0.090	−0.155**
	(0.075)	(0.058)	(0.074)
After × *Only UGC*	−0.126**	−0.057	−0.126**
	(0.061)	(0.043)	(0.060)
Observations	13,711	13,711	13,711
$\overline{R}^2$	0.892	0.877	0.894

Notes. The dependent variable is (log + 1) *Weekly Sales* in units in 2009. The term *Small UGC Share* indicates that more than 10% of the song-specific videos on the U.S. YouTube uploaded prior to April 1, 2009, were uploaded from an official account. The term *Official and UGC* indicates a share of official videos of between 0% and 10%. The term *Only UGC* indicates that no video was uploaded from an official account. The omitted category is *No Video*. Song and week fixed effects and the constant are not reported. Standard errors are in parentheses, clustered at the song/artist level.

*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

the new Vevo platform because Vevo operates within the YouTube platform in almost all countries except Germany. For 2013, we can observe not only unit sales in the physical and digital purchase channels (columns (1)–(3)) but also how often a particular song was streamed on free and premium *licensed closed streaming sites* (not including YouTube and Vevo; columns (4)–(6)).

We find point estimates remarkably similar to those in our analysis of the GEMA shock yet with the (expected) opposite sign. Looking at the difference-in-differences estimate in columns (1) and (4) of Table 9, the entry of Vevo increased total record sales by about 10% (90% CI, [2%, 18%]) and the number of weekly streams by about 16% (90% CI, [2%, 33%]). Although it is important to note that these coefficients are less precisely estimated than the coefficients reported in Table 5, perhaps because the entry of Vevo is a less clean experiment and because of the small number of physical copies sold in 2013, our Vevo results are not statistically different from the GEMA results because the absolute confidence bands overlap for all common coefficients. In the online appendix, we run a series of robustness checks that make us confident that the direction of our estimates reflects the causal effect of the Vevo entry. First, we show that the sales/streams of songs without Vevo videos and songs with Vevo videos follow similar trends before entry. Second, we show that the results hold under an alternative identification strategy (treating all songs released by Universal and Sony as affected). Third, we show that there are no significant effects when we run falsification exercises based on placebo timing or placebo geography. Fourth, we can rule out that the effect is driven by price changes that temporally coincide with the entry of Vevo.

5.4. Effect Heterogeneity

We now go beyond the average effect and look into effect heterogeneity. We allow for heterogeneity at the level of popularity in terms of sales, investigating the effect of the GEMA shock at various points of the sales distribution. We then allow for heterogeneity according to consumer awareness and the breadth of an artist's appeal. We estimate triple difference models by adding the additional interaction terms $\delta_1(After_t \times X_i)$ and $\delta_2(After_t \times Video_i \times X_i)$ to Equation (1), where $X_i \in \{Newcomer_i, Niche_i\}$. Under the null hypothesis $\delta_2 = 0$, we can directly test whether the effect of the GEMA shock differs across observations where $X_i = 0$ and $X_i = 1$. The total effect for observations where $X_i = 1$ is $\hat{\delta_1} + \hat{\delta_2}$.

Table 9. Average Effect of Vevo's Entry on Record Sales

	(1)	(2)	(3)	(4)	(5)	(6)
Variable	Total	Physical	Digital	Total	Free	Premium
After × *Video*	0.092**	0.021	0.087**	0.149*	0.133	0.143*
	(0.042)	(0.018)	(0.042)	(0.081)	(0.101)	(0.076)
Observations	14,576	14,576	14,576	14,576	14,576	14,576
$\overline{R^2}$	0.858	0.889	0.859	0.898	0.852	0.896

Notes. The dependent variable is (log + 1) *Weekly Sales/Streams* in units in 2013. The term *Video* indicates that at least one song-specific video was uploaded on the U.S. YouTube by Vevo. Song and week fixed effects and the constant are not reported. Standard errors are in parentheses, clustered at the song level.
$^*p < 0.10$; $^{**}p < 0.05$.

5.4.1. Overall Popularity. The results in Table 10 suggest that the average baseline effect does not differ significantly by overall popularity. We test for differences in the effect of the GEMA shock on sales across the distribution of the sales variable in a number of ways. We begin by exploring variation in the amount of time songs stay in the top 1,000 ranking to capture the notion that songs that stay in the top 1,000 ranking for longer are more popular. In column (1) of Table 10, we test for differences between songs that stay in the top 1,000 sales ranking throughout the observed period and those that drop out at some point. The effect of the GEMA shock on the latter, estimated as the coefficient *After* × *Video*, is −10% and therefore about four percentage points smaller than the coefficient in our baseline results in column (2) of Table 5. The 90% CI band ranges between −19% and −1%. We do not find a significant difference, as indicated in the coefficient of *After* × *Video* × *AlwaysInTop*1*k*. Adding the two, the total effect for songs that always stay in the top 1,000 ranking is −5%, but its 90% CI band overlaps substantially with that of the effect for songs that drop out of the top 1,000 list ([−15%, 5%]). Hence, we cannot interpret as evidence that more popular songs exhibit a less strong promotional effect from online music videos. In column (2) of Table 10, we go one step further and test whether the effect of the GEMA shock is different for even more popular songs. In the subset of songs that always remain in the top 1,000 list, we distinguish between those that never fall below rank 200 over the observed period and those that do.[30] Again, we find no significant difference. The effect for songs that exit the top 200 is −10% (90% CI, [−17%, −2%]), whereas the effect for those that stay within the top 200 is 2%, yet this estimate is very imprecise (90% CI, [−14%, 22%]). We continue to test whether songs at the top of the preexperiment sales distribution respond differently to the GEMA shock in column (3). Comparing songs in the 95th percentile with all others, we do not find a significant difference at the top of the preexperiment sales distribution. The implied effect for songs outside the 95th percentile is −12% (90% CI, [−19%, −4%]). The implied effect for songs in

Published in Information Systems Research on June 10, 2020 as DOI: 10.1287/isre.2019.0915.
This article has not been copyedited or formatted. The final version may differ from this version.

Table 10. Heterogeneity: Overall Popularity

	(1)	(2)	(3)	(4)	(5)	(6)
Variable	Time	Prerank	Presales	5th pctl.	95th pctl.	Weighted
After × *Video*	–0.110* (0.060)	–0.102** (0.052)	–0.125** (0.052)	–0.171** (0.074)	–0.207 (0.159)	–0.141* (0.080)
After × *AlwaysInTop1k*	–0.350*** (0.071)					
After × *Video* × *AlwaysInTop1k*	0.055 (0.088)					
After × *AlwaysInTop*200		–0.414*** (0.103)				
After × *Video* × *AlwaysInTop*200		0.124 (0.120)				
After × 95*th pctl.*			–0.547*** (0.133)			
After × *Video* × 95*th pctl.*			0.060 (0.154)			
Observations	13,711	2,799	13,711	13,711	13,711	13,711
$\overline{R^2}$	0.892	0.916	0.892			

Notes. The dependent variable is (log + 1) *Weekly Total Sales* in units. Song and week fixed effects are used in columns (1) and (5), columns (2)–(4) are quantile regressions with differenced-out song fixed effects and time dummies, and column (5) is a weighted regression using average preexperiment sales as weights. The term *Video* indicates that at least one song-specific video was uploaded on the U.S. YouTube prior to April 1, 2009. The constant is not reported. Standard errors are in parentheses, clustered at the song level. pctl., Percentile.
$^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

the 95th percentile is −6% but very imprecisely estimated (90% CI, [−26%, 19%]). Again, these results do not suggest that more popular songs exhibit a significantly weaker promotional effect from online music videos. In columns (4) and (5) of Table 10, we estimate quantile regressions for different percentiles of the dependent variable. Note that in these exercises, we look at pre- and postexperiment sales. The results suggest the effect of the GEMA shock is −16% at the 5th percentile of the total sales distribution (90% CI, [−25%, −5%]), and −19% at the 95th percentile (90% CI, [−37%, 6%]). Finally, in column (6), we report the results of a weighted regression, where we give observations with higher average sales in the preexperiment period a relatively larger weight. The point estimate is −13% (90% CI, [−24%, −1%]), still very similar to the baseline estimate in column (2) of Table 5.

These exercises do not support the notion that the average treatment effect in our setting affects different parts of the sales distribution differently. However, the above analysis did not consider that certain groups of observations may respond differently to the GEMA shock, which may affect sales of these groups but is subtle enough to not affect the overall popularity distribution. From a management and policy perspective, therefore, it is important to uncover this type of heterogeneity.

5.4.2. Consumer Awareness. Our measures of consumer awareness (defined in Section 4.4) are based on the idea that consumers are less aware of new artists than of established artists.[31]

Results in columns (1)–(3) of Table 11 show that newcomer artists are affected more strongly than established artists. The triple interaction term is negative in all columns and significant in columns (2) and (3), suggesting that total record sales of new artists decrease more when music videos are blocked on YouTube.

The point estimate in column (1) translates into a −12% effect for established artists (90% CI, [−19%, −5%]) and −50% for new artists (90% CI, [−75%, 2%]).[32] Column (2) suggests that total record sales decrease by 11% for established artists (90% CI, [−18%, −3%]), and by 54% for new artists (90% CI, [−75%, −14%]). The point estimates in column (3) of Table 11 are −11% for established artists (90% CI, [−18%, −3%]) and −46% for new artists (90% CI, [−50%, −41%]). These results are consistent with the notion that YouTube promotes artists who are less known to consumers. Although only the specification in column (3) shows that the differences between established and new artists are significant, it is important to note that the statistical precision in Table 11 is impacted by the fact that the number of observations that qualify as newcomers (in either definition) is relatively small.[33]

Table 11. Heterogeneity: Consumer Awareness

	(1)	(2)	(3)
Variable	*Two months*	*No album*	*First year*
After × *Video*	−0.133***	−0.114**	−0.114**
	(0.050)	(0.049)	(0.050)
After × *Newcomer*	0.668*	0.850***	0.637**
	(0.385)	(0.325)	(0.250)
After × *Video* × *Newcomer*	−0.553	−0.661*	−0.495*
	(0.430)	(0.384)	(0.287)
Observations	13,711	13,711	13,711
$\overline{R}^2$	0.892	0.892	0.892

Notes. The dependent variable is (log + 1) *Weekly Total Sales* in units. The term *Video* indicates that at least one song-specific video was uploaded on the U.S. YouTube prior to April 1, 2009. The term *Newcomer* is defined as *Two Months* in columns (1) and (3), as *No Album* in columns (2) and (4), and as *First Year* in columns (3) and (5). Song and week fixed effects and the constant are not reported. Standard errors are in parentheses, clustered at the song/artist level.
$^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

5.4.3 Breadth of Appeal. Our second set of results concerns heterogeneity across the inherent breadth of the artist's appeal. We do this by distinguishing between niche (narrow appeal) and mainstream (broad appeal) artists.[34] Because any empirical definition may be arbitrary, we again report results for three distinct measures (defined in Section 4.4). Results in Table 12 show that the triple interaction is positive and significant.

In column (1) of Table 12, the effect for nonniche artists is −33% (90% CI, [−46%, −16%]), whereas the effect for niche artists (defined as never having appeared in the U.S. charts) is −9% (90% CI, [−17%, −0.04%]). In column (2), the effect for non-German artists is −23% (90% CI, [−33%, −12%]), whereas for German-origin artists, we find a nonsignificant decrease of 8% (90% CI, [−17%, 3%]). The specification in column (3) distringuishes between pop/rock and other genres. The effect for pop/rock songs is −21% (90% CI, [−30%, −12%]), whereas for all other genres we find a nonsignificant increase of 1% (90% CI, [−10%, 12%]).

The 90% confidence bands overlap in the specifications in columns (1) and (2) of Table 12 and not in column (3). Hence, we can confirm only that total sales of artists in niche genres react differently to the GEMA shock than those of artists in mainstream genres for the genre-based definition of "niche."

6. Discussion

In this section, we first discuss how to interpret our findings in the context of music discovery and speculate about the underlying mechanisms before we assess the implications of the heterogeneous effects on aggregate sales. We then zoom out and provide a back-of-the-envelope calculation of the overall economic significance and welfare effects of the blocking policy in the context of the GEMA shock. Finally, we discuss the external validity and general implications of our study in relation to prior literature.

We find that the effect of the removal of online music videos on song sales is negative. Point estimates vary, but −10% to −5% is a conservative estimate of the short-run effects of removing access to music videos on sales of recorded music. This suggests that despite the fact that free interactive content could plausibly substitute for demand for paid content (as found in most of the piracy literature), free online music videos tend to complement sales of recorded music. The extent of UGC does not affect the positive effect of YouTube availability on music sales, a result established by directly measuring the share of UGC and further supported by the entry of Vevo. Our results further suggest that YouTube can play an important role in the discovery of music. Looking at artist subgroups, we find that new artists benefit more from content availability on YouTube (and consequently see a bigger drop in sales following the YouTube blackout). Conversely, niche artists do relatively better following the unavailability of YouTube content.

It is useful to think about these results in terms of the underlying process of music discovery. Because music is an experience good, consumers rely on sampling (Peitz and Waelbroeck 2006), word-of-mouth recommendations (Susarla et al. 2012, Lee et al. 2015), or automated recommender systems (Liu et al. 2014, Zhou et al. 2016) for their choice of which new musical content to consume. Underlying all of them is the notion that music discovery incurs search costs, and each of the three cues for consumers reduces search costs for new music. Consider now the discovery

Table 12. Heterogeneity: Breadth of Appeal

	(1)	(2)	(3)
Variable	Billboard	German	Genre
After × *Video*	−0.394***	−0.265***	−0.241***
	(0.134)	(0.081)	(0.066)
After × *Niche*	−0.310**	−0.213**	−0.061
	(0.131)	(0.083)	(0.079)
After × *Video* × *Niche*	0.300**	0.187*	0.249**
	(0.146)	(0.105)	(0.102)
Observations	13,711	13,711	13,711
$\overline{R}^2$	0.892	0.892	0.892

Notes. The dependent variable is (log + 1) *Weekly Total Sales* in units. The term *Video* indicates that at least one song-specific video was uploaded on the U.S. YouTube prior to April 1, 2009. The term *Niche* is defined as *Never U.S.* in column (1), as *German* in column (2), *Genre* in column (3). Song and week fixed effects and the constant are not reported. Standard errors are in parentheses, clustered at the song/artist level.
$^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

process that turns music videos on YouTube into music purchases. Exposure to free content on YouTube leads to higher sales under three conditions. First, "new" music has to be new to the individual, because otherwise the individual may have already purchased it in the past. Second, the new music has to meet the tastes of the listener. Finally, the sampled content must be a meaningful indicator of the quality of the song. Because the audio part of music videos tends to be either identical to or a close adaptation of what can be purchased in a (digital) record store, this condition holds in the case of YouTube. Discovery on YouTube may directly transform into digital purchases via click-to-buy links to retailers such as iTunes and Amazon (a mechanism that we cannot test in the absence of individual-level data). Other times, although there is a causal link between sampling on YouTube and music purchases, this may work more indirectly without leaving a trail in a consumer's clickstream.

What do our heterogeneous effects tell us about the way in which music discovery is affected by YouTube? First, the share of UGC plays no role in the effect of YouTube (un)availability. Conversion rates are not higher or lower if consumers hear the song accompanied by an official video compared with *UGC*, suggesting that the process of discovery does not need curated complements to songs. YouTube creates awareness of an original song (the function common to official and user-generated versions of the song), although the delivery channel is not considered a close substitute to actually purchasing the song.

We also find that new artists benefit relatively more from YouTube availability, as do mainstream artists. The former is in line with the awareness logic—listeners have not been exposed widely to these new artists, and exposure on YouTube helps listeners discover them. However, the result that mainstream artists benefit more than niche artists from YouTube availability suggests another function of YouTube: holding everything else fixed, artists in mainstream genres gain more from YouTube exposure. This suggests that the likelihood of a purchase after watching a YouTube video is higher for a mainstream artist because the general appeal of mainstream artists on YouTube is broader.

Apart from demand-side awareness and mass appeal, what supply-side mechanisms may lead to the disproportional effects we observe in the data? First, music videos may be systematically different for new and mainstream artists. Music videos have become a widespread complement to audio recordings only in recent years. This implies that new artists are more likely to have both more and better music videos than established artists.[35] Similarly, because of mainstream artists' greater expected appeal, the incentives to invest in video content are bigger, which would lead to more and perhaps better videos for mainstream songs. Another mechanism consistent with our heterogeneous effects is based on active platform management. YouTube's objectives will likely be a mix of dynamic and static goals. Maximizing views at any one point in time increases YouTube's opportunities for monetization in the short run. This speaks for promoting already popular content with wide appeal to increase the chances of repeat consumption. However, consumers need to be exposed to new content systematically to avoid them losing interest in the platform. This gives YouTube an incentive to recommend new artists.[36] In sum, both artists and labels (through better/more engaging video content for mainstream and new artists) and the platform itself (through recommender systems) may reinforce the process of music discovery driven by awareness and mass appeal of music. These supply-side actions are, however, complements to demand-side forces rather than substitutes.

6.1. Online Music Videos and Changes in the Distribution of Music Sales

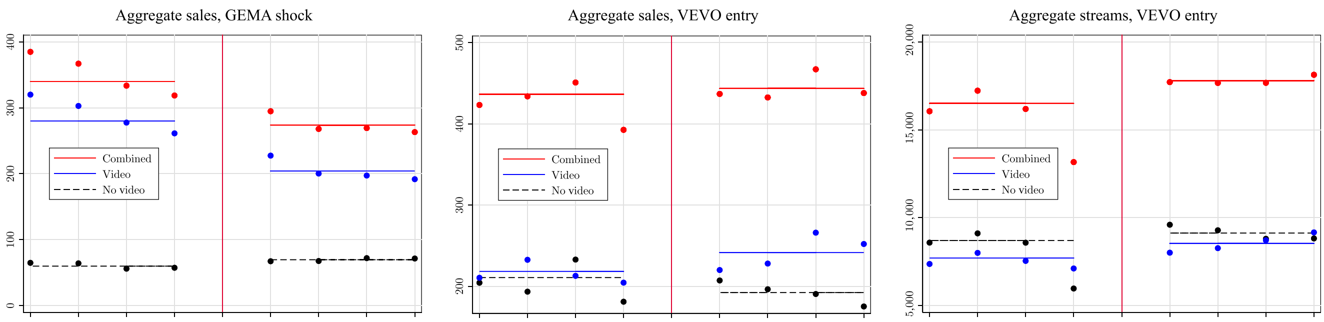
Our findings are consistent with the idea that the process of music discovery is affected by the availability of YouTube videos by raising potential consumers' awareness of new music, which especially benefits artists about whom consumers are initially relatively unaware and artists who have broad appeal and are therefore more likely to convert video views into record sales. This is in line with the work of Lee and Hosanagar (2019), who show that recommender systems in e-commerce tend to put weight on both extremes of the sales distribution; that is, popular products become more popular, but long-tail products increase sales too. The fact that different types of music benefit to different extents from exposure onYouTube can have implications for the overall sales distribution. The literature has long speculated on whether digital distribution platforms (and perhaps the recommender systems operating on them) help the most popular products become even more popular (Celma and Cano 2008, Fleder and Hosanagar 2009, Zhou et al. 2010, Oestreicher-Singer and Sundararajan 2012b) or whether they help niche products with relatively low sales (the long tail) reach a wider audience (Tucker and Zhang 2011, Oestreicher-Singer and Sundararajan 2012a). Related work finds evidence that streaming can also change the distribution of sales (Hiller 2016, Datta et al. 2018). Using individual-level consumption data from 900 consumers, Datta et al. (2018) show that users switching from owning (iTunes) to streaming (Spotify) listen to music from a more diverse set of artists and discover artists that are new to them. Similarly, the displacement effect of streaming in Hiller (2016) is weaker for less known albums.[37]

In the specific context of our short-run analysis of the removal of music videos because of the licensing dispute between GEMA and YouTube, a back-of-the-envelope calculation shows that newcomer artists suffer more than niche artists benefit relative to mainstream artists. A conservative reading of our results implies that sales of newcomer artists change by about $\hat{\delta}_{new} = -41\%$ and sales of niche-genre artists change by about $\hat{\delta}_{niche} = -10\%$ because of the GEMA shock. To understand the aggregate effect of this, we need to compare realized market shares with the counterfactual market shares that would have been realized had the GEMA shock not happened. We can estimate the counterfactual sales of newcomers and mainstream artists by dividing the observed total sales in the postshock period, that is, $\hat{T}_j^* = T_j/(1+\hat{\delta}_j)$, and then calculating their counterfactual market shares as $\hat{m}_j^* = \hat{T}_j^*/(\hat{T}_j^* + \hat{T}_{-j}^*)$. Realized market shares can be calculated accordingly, that is, $\hat{m}_j = \hat{T}_j/(\hat{T}_j + \hat{T}_{-j})$. Finally, we can express counterfactual market shares as percentages of realized market shares such that $\hat{\delta}_j^m = \hat{m}_j/\hat{m}_j^* - 1$. Average total weekly sales in the postshock period for newcomers (defined as in the first year of their career) and niche artists with music videos are $T_{new} = 24{,}301$ units and $T_{niche} = 22{,}384$ units, respectively. With this, we can calculate that the market share of newcomer artists decreases by 13%, whereas the market share of niche artists increases by 20%, as a result of the GEMA shock. Using more conservative estimates, that is, from the respective confidence bands of other measures of "newcomer" and "niche," leads to the same qualitative conclusions. On average, these calculations suggest that YouTube helps increase the sales of newcomer artists about two-thirds more than it helps increase the sales of mainstream artists.

After establishing the distributional changes across types of artists within the group of songs that have music videos, we now change the level of aggregation and investigate distributional changes across songs with and without music videos. So far, a valid concern would be that the availability of online music videos is a zero-sum game, that is, leading to changes in the composition of music sales without changing aggregate music sales. Furthermore, as in Liebowitz (2007), a fallacy of composition could lead to effects in the aggregate that are the opposite of the relationship on the individual song level.

We first analyze this issue at the song level. From visual inspection in Figure 2 and the corresponding figure in the online appendix, we see no indication that the trend in total sales/streams of songs without videos (our control group) changes after either the GEMA shock or the entry of Vevo. Accordingly, our results do not suggest that online music video platforms have externalities on the sales/streams of songs not available on these platforms. This is already indicative that the results we document are not merely driven by a transfer of demand from songs without videos to songs with videos. By looking at aggregate sales numbers, we can investigate this further. The plot of aggregate sales around the GEMA shock in the first panel of Figure 4 shows that also the aggregate sales trend of songs without videos does not change from the pre- to the postshock period. Aggregate sales of songs with videos, however, decline substantially. Accordingly, we also see a substantial decrease in total sales (songs with and without videos). Turning to aggregate sales in the time frame of the Vevo entry in the second panel of Figure 4, we see that songs without videos continue their negative sales trend, whereas aggregate sales of songs with videos increase. As a result, total record sales do not change by much. Finally, in the third panel of Figure 4, there is again no indication that the trend of aggregate

Figure 4. (Color online) Aggregate Effects of Online Music Video Platforms



Notes. In each panel, the vertical axis shows the aggregate of physical and digital sales in thousands of units. The horizontal axis shows weeks before and after the GEMA shock/Vevo entry. Dots indicate weekly aggregates, and lines indicate averages in the before and after periods. Aggregate sales of songs without music videos are shown in black. Aggregate sales of songs with music videos are shown in blue. Aggregate sales of all songs, with and without music videos, are shown in red.

streams of songs without videos changes with the entry of Vevo, especially if week t_{-1} is considered as an outlier. We do see, however, that aggregate streams of songs with videos increase. This leads to an increase in total streams (songs with and without videos).

Combining the insights from the disaggregated and aggregated analyses, we conclude that the experiments we study in this paper affect the composition of not only sales/streams but also aggregated sales/streams. However, note that the analysis of aggregated sales is of course less rigorous than the analysis using disaggregated data because the plots in Figure 4 are ultimately based on only eight data points.

6.2. Economic Significance and Welfare

Because it is difficult to attach a monetary value on the utility of free music consumption, it is difficult to derive an estimate of consumer surplus and therefore draw conclusions about overall welfare effects. However, we can calculate the average economic size of our estimates and derive estimates of total industry and artist surplus.

According to our price data, the average price for a song on a physical medium was €4.4 and that for a digital song was €1.1 after the GEMA shock. Multiplied by average total weekly sales units of songs with videos in the postshock period, this implies average total weekly revenues of about €360,000. With the most conservative estimate of a 6% reduction, this is 94% of the counterfactual revenues that would have been realized had the GEMA shock not happened. The monetary equivalent of 6% less sales than in the counterfactual world is thus about €23,000. Using the information that artists earn about 10 cents from a downloaded song and 13% of the physical revenue, we arrive at an estimate of a weekly decrease of about €2,500 in total artist income from record sales.[38]

It is difficult to say much about the loss in royalty income to artists because we do not have access to song-level data on the number of streams on YouTube prior to the GEMA shock. However, we do have data on the number of streams on services such as Spotify and Deezer for 2013. Hence, we can roughly calculate how the entry of Vevo has affected the surplus of the recorded music industry and the surplus of artists. Before launching in October 2013, Vevo signed a licensing deal with GEMA (see Cookson 2013). Hence, artists benefit from an increase in record sales and receive royalties collected by GEMA from Vevo.

Using our data on the average weekly sales and prices in Germany after the Vevo entry, the most conservative estimate of a 2% increase in total record sales translates to an increase in total income of about €4,500 and an increase in total artist income from record sales of about €450. According to GEMA's official royalty rates schedule for ad-funded streaming offerings (VR-OD-9, see GEMA 2019), the licensing fee for a highly interactive service such as Vevo is €0.00375 per stream. We do not have exact data on the number of streams on Vevo, but data from the United States can be helpful as an approximation. In 2013, there were 49.5 billion music streams, 22.4 billion (45%) from audio services and 27.1 billion (55%) from video services (see Ingham 2015). Our data on the average number of weekly (free and paid) audio streams in Germany do not include video streaming. Assuming that the ratio in the United States is the same in Germany, we can then calculate the hypothetical total royalties from video streaming as $(0.55/0.45) \times audio \times 0.00375$. The total number of audio streams of songs with Vevo videos in the postshock period is 7,894,000. This implies an increase in video streaming royalties worth roughly €730. We do not have data on Vevo's market share in the music video streaming market in Germany, but industry reports suggest that Vevo has had a market share of 2.8% in the overall online video market in 2016, which we can use as a conservative estimate.[39] We thus arrive at a lower bound of the average weekly increase in video streaming royalties of about €300. The average weekly increase in audio streaming royalties is about €245. Our estimate of the total increase in artist surplus is therefore some €700 per week. However, this includes only artists who are represented by Vevo (~70% of the market).

Now we can calculate the total lost artist surplus from the unavailability of online music videos—the sum of lower record sales and foregone royalty income from video streaming. Assuming that the number of video streams is stable over time, foregone royalties of all artists amount to about €400 per week. This implies that the promotional externalities of online music video (€2,500 total artist surplus) can offset forgone royalty income by a factor of six. In the 235 weeks between the GEMA blocking and Vevo entry, our most conservative estimate is a total loss in artist welfare of about €0.6 million and a loss in industry revenues of €5.4 million. Note that both numbers do not include potential advertising revenues from YouTube.

6.3. External Validity and Contribution

6.3.1. Prior Work on the Effect of YouTube on Recorded Music Sales. At face value, our baseline result seems in direct contrast to that of Hiller (2016), who concludes that a royalty dispute that led to the takedown of Warner Music content on YouTube is related to an increase in sales of albums released by Warner. However, a careful analysis of the differences and commonalities in the two studies reveals that some of Hiller's (2016) results are line with our findings.

We first discuss key differences between Hiller's (2016) approach and ours and test some of the resulting empirical implications. We then try to replicate Hiller's (2016) approach as closely as possible in our setting and discuss alternative interpretations of Hiller's (2016) results and how they resonate with our findings.

Hiller's (2016) sample comprises the 200 best-selling albums in the U.S. market. He does not link sales and corresponding video availability at the individual song level but compares album sales of Warner artists with album sales of non-Warner artists. The distinction between songs and albums can have important implications. Because albums are bundles consisting of individual songs, YouTube availability, by allowing consumers to sample, may lead consumers to substitute purchases of the entire album in favor of purchases of an individual song that is promoted by a music video. In the absence of the sampling mechanism on YouTube, sales of other songs that are not promoted by a music video (or the entire album bundle) could therefore increase. We can at least partly test this hypothesis.

In column (1) of Table 13, we investigate how the removal of video(s) for a specific song affects sales of other songs (that do not have videos) on the same release. Releases with multiple songs are typically extended singles, minialbums (EPs), and, in the digital channel also, entire albums. We find a positive and significant coefficient.[40] The point estimate is 22% (90% CI, [7%, 38%]). Hence, this suggests that Hiller's (2016) results can be driven by unbundling and substitutive effects within releases.

In columns (2)–(4) of Table 13, we test whether we can replicate Hiller's (2016) results when we aggregate sales at the release level. Although this is not the same as actual album sales—especially not in the physical channel—it can serve as a proxy for album sales. The first identification strategy is to compare releases that have at least one song with a video on YouTube with those that do not. The effect in column (2) is −16% (90% CI, [−26%, −4%]), which is similar to our baseline effect. This implies that in our setting, the song-level substitution effects that we find in column (1) are not strong enough to outweigh the release-level promotional effects of music videos.[41]. In column (3), we follow Hiller (2016) and restrict our sample to the sum of sales within a release while songs are in the top 200 list. This decreases precision of the estimates, leading to an effect of −12% that is not significant (90% CI, [−39%, 24%]).

Finally, we adjust the identification strategy to better align with Hiller (2016). The equivalent to Hiller's (2016) approach of comparing sales of Warner artists with sales of non-Warner artists in our setting is to compare sales in Germany with sales in a comparable country, say, Austria. The caveat is that we can run this analysis only for sales in the physical channel because sales data for the digital channel are not available for Austria. In column (4) of Table 13, we find a negative yet imprecisely estimated coefficient for *After* × *Germany*, with a point estimate of−2% (90% CI, [−6%, 1%]). Much as in Section 5.2.2, the less precise identification strategy leads to smaller point estimates, but the sign of the effect remains. Finally, our closest replication of Hiller's (2016) approach, in column (5) of of Table 13, considers only the sum of sales within a release while songs are in the top 200 list. We find a positive coefficient for *After* × *Germany* that implies an effect of 31% (90% CI, [−0.2%, 73%]).

Table 13. Aggregate (Release-Level) Effects of the GEMA Shock

	(1)	(2)	(3)	(4)	(5)
Variable	*OtherSongs*	*Release*	*ReleaseTop*	*Release*	*ReleaseTop*
After × *OtherVideo*	0.197**				
	(0.077)				
After × *AtLeastOneVideo*		−0.173**	−0.126		
		(0.078)	(0.223)		
Germany				0.435***	2.156***
				(0.043)	(0.187)
After × *Germany*				−0.024	0.272
				(0.021)	(0.166)
Observations	9,919	6,841	1,180	25,102	2,259
$\overline{R^2}$	0.902	0.894	0.939	0.031	0.282

Notes. The dependent variable is (log + 1) *Weekly Total Sales* in units, 2009. Column (1) uses physical and digital sales of individual songs. Column (2) uses physical and digital sales of the release bundle. Column (3) uses physical and digital sales of the release bundle only when songs are part of the top 200 list. Column (4) uses physical sales of the release bundle. Column (5) uses physical sales of the release bundle only when songs are part of the top 200 list. The variable *AtLeastOneVideo* (*OtherVideo*) indicates at least one (other) song on the same release has a video on the U.S. YouTube uploaded prior to April 1, 2009. Song and week fixed effects are used in column (1), release and week fixed effects are used in columns (2)–(4), and the constant is not reported. Standard errors are in parentheses, clustered at the song level.

$^{**}p < 0.05$; $^{***}p < 0.01$.

This is very similar to the effect reported in column (1) of table 4 in Hiller (2016). The stark disparity between the results in columns (4) and (5) suggests that the empirical approach of Hiller (2016), a combination of sample restrictions and measurement error in the definition of treatment and control group, may be responsible for the displacement effect that he finds. Furthermore, it seems likely that Hiller's (2016) finding is mainly driven by the most popular albums.

Indeed, this can be seen directly by looking more closely at Hiller's (2016) results. He shows that albums that have a very successful debut face more displacement from YouTube videos, whereas the effect on lower debuting albums may be moderated by a promotional effect. In fact, the estimated interaction effect is so strong that the relationship is reversed, and the promotional effect dominates for many albums in the sample. The results in table 5 in Hiller (2016) suggest that *Debutrank* moderates the average effect such that the sign changes with *Debutrank* greater than values between 15.64 and 44.41 (depending on the specification, $x = -\beta_{Warnereffect}/\beta_{Warnerdebutrank}$), which is much lower than the average *Debutrank* of Warner albums in his sample (79.4). A valid alternative interpretation of Hiller's (2016) results, therefore, is that most albums' sales benefit from the added awareness generated through YouTube, and only the most popular albums (those with very small debut ranks) experience net *displacement* rather than *promotion*. As we have shown in Section 5.4, we do not find heterogeneity across the overall sales distribution (as Hiller (2016) does), but we do find that the promotional effect is driven by new artists and by artists in nonniche genres (which tend to rank slightly better, that is, are more popular, when they enter the top 1,000).

All this leads us to conclude that the differences between our results and Hiller's (2016) findings are largely due to the of level of aggregation, sample restrictions, and the choice of identification strategy. Hiller (2016) provides a partial view on the phenomenon at hand, whereas our analysis covers a broader range of the popular appeal of music content through the deeper analysis afforded by our broader and more detailed data and our more fine-grained identification strategy. Our findings show important heterogeneity, suggesting that the average effect hides sizable differences in the strength of one of the counteracting forces, the role of YouTube availability on music discovery. Thus, we offer a more detailed unpacking of the process of music discovery than Hiller (2016). This lets us draw more comprehensive policy implications.

6.3.2. Policy Implications. Although our study is clearly limited in its scope, being a short-run measurement exercise that makes use of natural experiments that happened in specific contexts, it can still be useful to think about the policy implications that arise from our results. First, the policy of some countries to mandate a share of broadcast music to be local (or in the local language) may be affected by the role of interactive and global platforms such as YouTube. Second, U.S. copyright law currently stipulates licensing fees for interactive digital platforms but none for radio broadcasting based on assumptions about the degree of substitution between (free) digital and paid content (Liebowitz 2007, Lenard and White 2015). Finally, from the perspective of the German right holders' association, the differential effects on domestic versus international artists may imply that it was acting to maximize their direct stakeholders' revenues. We elaborate on each of these domains next.

Several countries (e.g., France and Portugal) consider local music a cultural good in danger of being overrun and eventually replaced by international, often English-language content. This has led to quasi-protectionist policies stipulating a minimum percentage (often 40% or 50%) of local content on radio and television broadcasts (Hervas-Drane and Noam 2017). Our results suggest that although YouTube does not displace music sales (which would provide a direct, unregulated channel for international music to enter the local market), music discovery via YouTube favors mainstream (non-German in our case) artists substantially more than local ones.[42] This implies that policies to protect local (national) creative music industries are less effective in the presence of digital platforms for music discovery.

In the United States, the rules of compulsory licensing imply that the music industry receives licensing fees from interactive digital platforms, whereas radio broadcasts are traditionally exempt from licensing. It is interesting to consider the historical reasoning behind the exemption: radio is a unidirectional medium on which users cannot choose particular songs, and it was heavily influenced by the music industry, which offered payment or other inducements to radio stations to play particular songs (Coase 1979). The perception was that radio acted as a promotional channel for actual purchases, relieving broadcasters of the need to compensate artists for playing their music because the compensation would come in the form of higher sales.[43] Conversely, digital platforms are on-demand and users can (at least to some extent) choose the songs they want to hear, which in principle allows for more substitution. Our findings suggest that digital platforms also fulfill an important promotional function, at least for some types of music. If replicated in other settings, such evidence might justify extending the license fee exemption to plays on digital platforms, and for firms, this result offers a case for differential online royalties, with new music especially prone to benefiting from exposure on digital platforms.

Finally, it is also worth revisiting the outcome for one of the instigators of the initial dispute, GEMA. GEMA's members are (virtually all) German artists, with international artists associated only by virtue of their membership in other national rights holders' associations. Our findings show that German artists suffered relatively less from the YouTube blackout and gained market share as a result.[44] Therefore, the removal of videos on YouTube may have (deliberately or not) resulted in the main constituency of GEMA increasing its market position vis-à-vis its international competitors.

7. Conclusions

In this paper, we exploit two natural experiments in the German market for online music videos to identify the effect of free sampling on sales of recorded music. The first experiment lets us identify the effect of removing access to online music videos on YouTube (in April 2009), whereas in the second, we identify the effect of making official music videos available on the proprietary platform Vevo (in October 2013). Our analysis is based on a rich data set that combines sales data that cover a large fraction of all music sales with song-level information on music video availability.

We find robust evidence that online videos are complementary to record sales. We believe that our findings carry some important general implications. Three results especially suggest to us that the promotional effect we robustly identify is not driven by the fact that YouTube is an open platform but rather by the fact that YouTube offers differentiated content—music *video* rather than music. First, we show that YouTube's effect on music sales is driven not only by the availability of official music videos but also by UGC on YouTube. Second, our results are very similar when we look at the entry of the closed platform Vevo. Third, we show that the promotional effect prevails when we look at the number of plays on audio streaming platforms, which should be a very close substitute for music video streaming.

Although we do not find the effect to differ based on overall popularity in terms of sales, we show that sales dynamics of newcomer artists are affected much more by the (non)availability of online music videos than those of established artists. We also find that YouTube disproportionally favors music genres with greater mass appeal. We discuss several possible mechanisms consistent with the observed dynamics on the demand side and on the supply side. In contrast to the related literature, we study the effect of online music videos on the sales of songs and therefore cannot rule out that the effects we show do not hold regarding bundled sales of songs, for example, albums.

In reference to a song by *The Buggles*—which happens to be the first music video shown on MTV in 1981—we conclude that our study does not provide much evidence that "video killed the radio star."[45] If anything, we find the opposite. Whereas it is straightforward to conclude that free consumption increases consumer surplus, conclusions about overall welfare are more difficult to reach. We calculate that positive externalities of access to an open platform such as YouTube can offset forgone royalty income (if YouTube did not pay any royalty fees to artists) by a factor of six. Furthermore, we show that the distributional effects are substantial. We calculate that YouTube helps to increase the sales of newcomer artists by about two-thirds more than it helps to increase the sales of mainstream artists.

Acknowledgments

The authors thank the senior editor; the associate editor; three anonymous reviewers; and Joel Waldfogel, Lisa George, Christiaan Hogendorn, Julian Wright, Ulrich Kaiser, Miguel Godinho de Matos, Rahul Telang, Danielle Li, Imke Reimers, Johannes Koenen, and Leon Zucchini for their valuable comments and ideas. They also thank Alexander Wolf and Tina Funk (executives at GEMA and Vevo, respectively) for insightful discussions. This paper benefited from feedback from conference participants at the National Bureau of Economic Research Summer Institute (2014), the International Industrial Organization Conference (2014), the Searle Internet Conference (2014), Verein für Socialpolitik Industrieökonomischer Ausschuss (2014), the Toulouse Digital Economy Workshop (2017), the Florence Seminar on Media, INFORMS, the Media Economics Workshop (2014), the ICT Conference Munich (2013), and the ZEW ICT Conference (2013); and participants at seminar presentations at the University of Zurich, Ludwig Maximilian University, the Toulouse School of Economics, ParisTech, USC Marshall, the Rotterdam School of Management, Católica-Lisbon, Pompeu Fabra University, Georgia Tech, the Tilburg Law and Economics Center, and the Institute for Prospective Technological Studies in Seville.

Endnotes

[1] See, for example, Hui and Png (2003), Rob and Waldfogel (2006), Zentner (2006), Rob and Waldfogel (2007), Oberholzer-Gee and Strumpf (2007), and the survey in Liebowitz (2016), although recent evidence in Peukert et al. (2017) suggests that there is significant heterogeneity across content types.

[2] See O'Brien (2009). More than seven years later, an agreement was reached; see BBC News (2016).

[3] See http://apps.opendatacity.de/gema-vs-youtube/en.

[4] See Section 6.3.1 for a discussion of a related paper (Hiller 2016) that studies the relationship of YouTube and album sales.

[5] See Ingham (2016). YouTube's influence on the music industry eventually grew so large that *Billboard* started incorporating YouTube data into their rankings in February 2013 (see *Billboard* 2013).

[6] The case of Taylor Swift withdrawing her songs from Spotify but not YouTube is a case in point (see Solon 2016).

[7] Examples for international counterparts are BMI, ASCAP, and SESAC in the United States, PRS in the United Kingdom, SACEM in France, and SGAE in Spain.

[8] See Schuetze (2012). In 2009, GEMA had 64,534 members and distributed €713 million in royalties.

[9] There is no evidence that YouTube systematically warned content owners in Germany before blocking videos.

[10] Specific legal issues have made it complicated to reach an agreement between GEMA and YouTube. According to Rolf Budde, member of the GEMA advisory board, YouTube insists on a nondisclosure agreement (see https://www.youtube.com/watch?v=Hh3Ks4Kxvtk). However, GEMA is required by law to publish the exact royalty payment schemes. Reportedly, because of this deadlocked situation, the involved parties consulted the arbitration board of the German Patent and Trademark Office for mediation in January 2013, and an agreement was finally reached in November 2016 (see BBC News 2016).

[11] After careful research, we could find only anecdotal evidence of one band opting out of GEMA. Videos on the official YouTube channel of the successful German rock band Die Ärzte were accessible in Germany (see *Spiegel Online* 2012a). It is not clear whether the band opted out of GEMA. When we asked the band's management for a statement, it declined to comment on the issue.

[12] See https://drive.google.com/open?id=0Bxe11iVXrXgsM0NIQm1OV1QxSVk.

[13] Furthermore, there is no evidence that consumers switched to other video platforms in the short run.

[14] For example, Sven Regener, singer of Element of Crime, says (referring to YouTube), "a business model based on people who produce the content not getting any money is not a business model, it's crap" (see http://www.br.de/radio/bayern2/sendungen/zuendfunk/regener_interview100.html). The popular electro/hip-hop band Deichkind posted a raging comment on its Facebook page after finding out that its newly uploaded music video was being blocked (see Spiegel Online 2012b).

[15] Market share data according to Nielsen SoundScan for the United States (see http://www.statista.com/statistics/317632/market-share-record-companies-label-ownership-usa/).

[16] During the observation period, physical record stores drastically reduced shelf space for singles and focused on physical albums (see Wallop 2008).

[17] See http://www.musikindustrie.de/umsatz/.

[18] Python code can be found on the authors' GitHub page: https://github.com/cpeukert/.

[19] This query was performed on April 15, 2015. In estimations not reported here but available on request, we get very similar results if we use information obtained from the Austrian version of YouTube.

[20] George and Peukert (2014) define official videos based on whether the video title includes the word "official," whereas our definition is more accurate in that it is based on the account that uploaded the video. See below for details.

[21] Furthermore, in an analysis available on request, we show that songs with videos tend to be newer than songs without videos, songs with more recent videos tend to have higher sales, and newer songs (that are as old as YouTube itself, which started in April 2015, or newer) more or less immediately have videos, whereas older songs show no clear pattern.

[22] According to what Tina Funk, general manager for Germany at Vevo, told us, this mostly holds true, but Vevo Germany also has some exclusive content, especially in the first weeks of its launch.

[23] All 247 artists listed under Vevo's main YouTube account have "Vevo" in their YouTube username; see https://www.youtube.com/user/Vevo/channels?view=56. We also manually checked all artists in our sample to make sure that we did not miss a Vevo account that does not follow this convention.

[24] For the example *Niche: Never U.S.*, $P(B|A) = 0.821$ is the probability that a song is from niche artist if it does not have a video, $P(B) = 0.723$ is the probability that a song is from a niche artist, and $P(A) = 1 - 0.548 = 0.452$ is the probability that a song does not have a video. Using Bayes' theorem, $P(A|B) = (0.821 \times 0.452)/0.723 = 0.513$ gives the probability that a song does not have a video if it is from a niche artist.

[25] In results available on request, we show that the estimated effect of the GEMA shock is similar when we do not include fixed effects. Furthermore, the estimated effect is consistent if we look at sales ranks, rather than units, as the dependent variable.

[26] Here, and in what follows below, we calculate percentage effects as $(exp(-0.142) - 1) \times 100 = -13.24$.

[27] If artists quickly adjust their digital strategy and drop out of GEMA or if consumers quickly discover technical measures to circumvent the blocking (e.g., a virtual private network), we expect a positive coefficient of *After* × *Video* in the second placebo test.

[28] One may be concerned that the standard errors are large because the sample size is much smaller than in the baseline specifications of Table 5. To make a fair comparison, we should therefore also estimate the "true" effect on a smaller sample. In results available on request, we show that the coefficient of *After* × *Video* estimated on a sample covering a window of ±2 weeks around the true date is −0.133 (standard error, 0.051), with a 90% CI ranging from −0.233 to −0.033.

[29] In results available on request, we show that the coefficients remain similar (qualitatively the same) if we use different thresholds, for example, the sample mean of 6.4%.

[30] We choose the cutoff of 200 to be consistent with Hiller (2016). See the discussion in Section 6.3.1.

[31] In the online appendix, we show that the parallel trends assumption also holds for newcomer artists as a subset of the treatment group.

[32] Percentage effects are calculated as $(\exp(-0.133) - 1) \times 100 = -12.45$ and $(\exp(-0.133 - 0.553) - 1) \times 100 = -49.64$, respectively.

[33] This is, of course, a reflection of the empirical distribution: the number of new artists in the music market is almost mechanically smaller than the number of established artists.

[34] In the online appendix, we show that the parallel trends assumption also holds for niche artists as a subset of the treatment group.

[35] Note that established artists in our data set include ones that started their musical careers before MTV and music videos were "invented" and widely adopted as marketing tools.

[36] This is highly consistent with the actual goals of YouTube's video recommendation system as practiced during our time period, as the following quote (from a paper by researchers at Google that describes YouTube's recommendation algorithm) shows: "We want recommendations to be reasonably recent and fresh, as well as diverse and relevant to the user's recent actions" (Davidson et al. 2010, p. 294).

[37] See Section 6.3.1 for a detailed discussion of Hiller (2016) in relation to our findings.

[38] See http://www.informationisbeautiful.net/2010/how-much-do-music-artists-earn-online and Donovan (2013).

[39] See https://de.statista.com/statistik/daten/studie/209329/umfrage/fuehrende-videoportale-in-deutschland-nach-nutzeranteil.

[40] Note that this does not affect the identification strategy in the main analysis; see the tests in Section 5.1.

[41] Note that this reconfirms our findings in Section 6.1 regarding the fallacy of composition hypothesis.

[42] The results discussed in Section 5.4.3 imply that removing videos of non-German artists leads to a significantly lower sales (point estimate, −23%). The sales effect of removing videos of German artists is about three times smaller and not significant (point estimate, −8%).

[43] The results in Liebowitz (2007) suggest that the opposite is true. He shows that radio is more of a substitute for the purchase of sound recordings than it is a complement.

[44] We can estimate market share changes in a similar fashion, as described in Section 6.1. Our results imply that the market share of non-German artists with videos decreased by 6% (average total realized sales are 130,898 units), whereas the market share of German artists with videos increased by 13% (average total realized sales are 72,864 units), as a result the removal of music videos on YouTube.

[45] According to music scholar Timothy Warner, the song's lyrics are "concerned with the adverse effect of technological change" (Warner 2003, p. 47) on artists that used to be commercially successful (in the "Golden Age of Radio," Warner 2003, p. 44). This relates to our study, which is essentially an exercise of measuring the substitutionary or complementary effects of different distribution technologies.

References

Adermon A, Liang CY (2014) Piracy and music sales: The effects of an anti-piracy law. *J. Econom. Behav. Organ.* 105(September): 90–106.

Aguiar L, Waldfogel J (2016) Even the loser get lucky sometimes: New products and the evolution of music quality since Napster. *Inform. Econom. Policy* 34(March):1–15.

Aguiar L, Waldfogel J (2018a) Quality predictability and the welfare benefits from new products: Evidence from the digitization of recorded music. *J. Political Econom.* 126(2):492–524.

Aguiar L, Waldfogel J (2018b) Streaming reaches flood stage: Does spotify stimulate or depress music sales? *Internat. J. Indust. Organ.* 57(March):278–307.

Anderson C (2006) *The Long Tail: How Endless Choice is Creating Unlimited Demand* (Random House, New York).

Autor DH (2003) Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing. *J. Labor Econom.* 21(1):1–42.

BBC News (2016) YouTube's seven-year stand-off ends. (November 1), http://www.bbc.com/news/technology-37839038.

Bertrand M, Duflo E, Mullainathan S (2004) How much should we trust differences-in-differences estimates? *Quart. J. Econom.* 119(1): 249–275.

Bhattacharjee S, Gopal R, Lertwachara K, Marsden J, Telang R (2007) The effect of digital sharing technologies on music markets: A survival analysis of albums on ranking charts. *Management Sci.* 53(9):1359–1374.

Billboard (2013) Billboard charts add YouTube views. (February 21), http://www.billboard.com/biz/articles/news/1549766/billboard-charts-add-youtube-views.

Brynjolfsson E, Hu YJ, Simester D (2011) Goodbye Pareto principle, hello long tail: The effect of search costs on the concentration of product sales. *Management Sci.* 57(8):1373–1386.

Brynjolfsson E, Hu YJ, Smith MD (2010) Long tails vs. superstars: The effect of information technology on product variety and sales concentration patterns. *Inform. Systems Res.* 21(4):736–747.

Cai H, Chen Y, Fang H (2009) Observational learning: Evidence from a randomized natural field experiment. *Amer. Econom. Rev.* 99(3): 864–882.

Celma Ò, Cano P (2008) From hits to niches?: Or how popular artists can bias music recommendation and discovery. *Proc. 2nd KDD Workshop Large-Scale Recommender Systems Netflix Prize Competition* (ACM, New York), 5.

Coase RH (1979) Payola in radio and television broadcasting. *J. Law Econom.* 22(2):269–328.

Comscore (2013) Comscore releases December 2012 U.S. online video rankings. Report, Comscore, Reston, VA.

Cookson R (2013) Vevo moves into Germany without YouTube. *Financial Times* (September 1), https://www.ft.com/content/5bb2092e-117e-11e3-a14c-00144feabdc0.

Danaher B, Dhanasobhon S, Smith MD, Telang R (2010) Converting pirates without cannibalizing purchasers: The impact of digital distribution on physical sales and internet piracy. *Marketing Sci.* 29(6):1138–1151.

Danaher B, Smith MD, Telang R, Chen S (2014) The effect of graduated response anti-piracy laws on music sales: Evidence from an event study in France. *J. Indust. Econom.* 62(3):541–553.

Datta H, Knox G, Bronnenberg B (2018) Changing their tune: How consumers' adoption of online streaming affects music consumption and discovery. *Marketing Sci.* 37(1):5–21.

Davidson J, Liebald B, Liu J, Nandy P, Van Vleet T, Gargi U, Gupta S, et al. (2010) The YouTube video recommendation system. *Proc. 4th ACM Conf. Recommender Systems* (ACM, New York), 293–296.

Donovan N (2013) If CDs cost £8 where does the money go? *BBC News Magazine* (August 26), http://www.bbc.com/news/magazine-23840744.

Ferreira F, Waldfogel J (2013) Pop internationalism: Has half a century of world music trade displaced local culture? *Econom. J.* 123(569):634–664.

Fleder D, Hosanagar K (2009) Blockbuster culture's next rise or fall: The impact of recommender systems on sales diversity. *Management Sci.* 55(5):697–712.

Galasso A, Schankerman M (2015) Patents and cumulative innovation: Causal evidence from the courts. *Quart. J. Econom.* 130(1): 317–369.

Gambardella A, Raasch C, von Hippel E (2017) The user innovation paradigm: Impacts on markets and welfare. *Management Sci.* 63(5):1450–1468.

GEMA (2019) Music and music video streaming. Accessed March 6, 2020, https://www.gema.de/fileadmin/user_upload/Musiknutzer/Tarife/Tarife_VRA/tarif_vr_od9_e.pdf.

George LM, Peukert C (2014) Youtube decade: Cultural convergence in recorded music. Working Paper 14-11, Networks, Electronic Commerce, and Telecommunications Institute, New York.

Godinho de Matos M, Ferreira P, Smith MD, Telang R (2016) Culling the herd: Using real-world randomized experiments to measure social bias with known costly goods. *Management Sci.* 62(9): 2563–2580.

Hervas-Drane A (2015) Recommended for you: The effect of word of mouth on sales concentration. *Internat. J. Res. Marketing* 32(2): 207–218.

Hervas-Drane A, Noam E (2017) Peer-to-peer file sharing and cultural trade protectionism. *Inform. Econom. Policy* 41(12): 15–27.

Hiller RS (2016) Sales displacement and streaming music: Evidence from YouTube. *Inform. Econom. Policy* 34(March):16–26.

Hinz O, Eckert J, Skiera B (2011) Drivers of the long tail phenomenon: an empirical analysis. *J. Management Inform. Systems* 27(4):43–70.

Hosanagar K, Fleder D, Lee D, Buja A (2014) Will the global village fracture into tribes? recommender systems and their effects on consumer fragmentation. *Management Sci.* 60(4):805–823.

Hui KL, Png I (2003) Piracy and the legitimate demand for recorded music. *B.E. J. Econom. Anal. Policy* 2(1):1–24.

Ingham T (2015) YouTube is the no.1 music streaming platform—And getting bigger. *Music Bus. Worldwide* (July 6), https://www.musicbusinessworldwide.com/youtube-is-the-no-1-music-streaming-platform-and-getting-bigger/.

Ingham T (2016) YouTube's dominance takes a beating from Spotify, Apple Music and co. *Music Bus. Worldwide* (July 5), https://www.musicbusinessworldwide.com/youtubes-dominance-takes-a-beating-from-spotify-apple-music-and-co/.

International Federation of the Phonographic Industry (2016) Music consumer insight report 2016. Report, Ipsos and International Federation of the Phonographic Industry, London. Accessed March 6, 2020, https://drive.google.com/open?id=0Bxe11iVXrXgsYXcxNUxUMFBOWHc.

King D (2008) Making money on YouTube with Content ID. *Google Official Blog* (August 27). Accessed May 2017, https://googleblog.blogspot.com/2008/08/making-money-on-youtube-with-content-id.html.

Kretschmer M (2005) Artists' earnings and copyright: A review of British and German music industry data in the context of digital technologies. *First Monday* 10(1).

Kumar A, Tan Y (2015) The demand effects of joint product advertising in online videos. *Management Sci.* 61(8):1921–1937.

Lee D, Hosanagar K (2019) How do recommender systems affect sales diversity? A cross-category investigation via randomized field experiment. *Inform. Systems Res.* 30(1):239–259.

Lee YJ, Hosanagar K, Tan Y (2015) Do I follow my friends or the crowd? Information cascades in online movie ratings. *Management Sci.* 61(9):2241–2258.

Lenard T. M, &White L. J (2016). Moving music licensing into the digital era: More competition and less regulation. *UCLA Entertainment Law Rev.* 23(1):133–156.

Liebowitz SJ (2007) Don't play it again Sam: Radio play, record sales, and property rights. Preprint, submitted January 11, http://dx.doi.org/10.2139/ssrn.956527.

Liebowitz SJ (2016) How much of the decline in sound recording sales is due to file-sharing? *J. Cult. Econom.* 40(1):13–28.

Liikkanen LA, Salovaara A (2015) Music on YouTube: User engagement with traditional, user-appropriated and derivative videos. *Comput. Human Behav.* 50(September):108–124.

Liu JH, Zhou T, Zhang ZK, Yang Z, Liu C, Li WM (2014) Promoting cold-start items in recommender systems. *PLoS One* 9(12): e113457.

Meusers R (2012a) Netzwelt-Ticker: Warum das neue Ärzte-Album komplett auf YouTube läuft. *Spiegel Online* (April 18), http://www.spiegel.de/netzwelt/web/netzwelt-ticker-warum-das-neuex-aerzte-album-komplett-auf-youtube-laeuft-a-828244.html.

Meusers R (2012b) Deichkind zum Gema-Streit: "Ihr seid Evolutionsbremsen." *Spiegel Online* (December 3), http://www.spiegel.de/netzwelt/web/deichkind-zum-gema-streit-ihr-seid-evolutionsbremsen-a-820703.html.

Nguyen GD, Dejean S, Moreau F (2014) On the complementarity between online and offline music consumption: The case of free streaming. *J. Cultural Econom.* 38(4):315–330.

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

O'Brien K (2009) Royalty dispute stops music videos in Germany. *New York Times* (April 2), http://www.nytimes.com/2009/04/03/technology/internet/03youtube.html.

Oestreicher-Singer G, Sundararajan A (2012a) Recommendation networks and the long tail of electronic commerce. *MIS Quart.* 36(1):65–84.

Oestreicher-Singer G, Sundararajan A (2012b) The visible hand? Demand effects of recommendation networks in electronic markets. *Management Sci.* 58(11):1963–1981.

Papies D, Eggers F, Wlömert N (2011) Music for free? How free ad-funded downloads affect consumer choice. *J. Acad. Marketing Sci.* 39(5):777–794.

Peitz M, Waelbroeck P (2006) Why the music industry may gain from free downloading—The role of sampling. *Internat. J. Indust. Organ.* 24(5):907–913.

Peukert C, Claussen J, Kretschmer T (2017) Piracy and box office movie revenues: Evidence from Megaupload. *Internat. J. Indust. Organ.* 52(May):188–215.

Piolatto A, Schuett F (2012) Music piracy: A case of "the rich get richer and the poor get poorer." *Inform. Econom. Policy* 24(1):30–39.

Poort J, Weda J (2015) Elvis is returning to the building: Understanding a decline in unauthorized file sharing. *J. Media Econom.* 28(2):63–83.

Raffo J, Lhuillery S (2009) How to play the names game: Patent retrieval comparing different heuristics. *Res. Policy* 38(10):1617–1627.

Rob R, Waldfogel J (2006) Piracy on the high C's: Music downloading, sales displacement, and social welfare in a sample of college students. *J. Law Econom.* 49(4):29–62.

Rob R, Waldfogel J (2007) Piracy on the silver screen. *J. Indust. Econom.* 55(3):379–395.

Salganik MJ, Dodds PS, Watts DJ (2006) Experimental study of inequality and unpredictability in an artificial cultural market. *Science* 311(5762):854–856.

Schuetze B (2012) The GEMA—Presumption and the burden of non liquet (Germany). *Kluwer Copyright Blog* (October 1), http://kluwercopyrightblog.com/2012/10/01/the-gema-presumption-and-the-burden-of-non-liquet-germany/.

Solon O (2016) Taylor Swift is taking on YouTube, and it won't be an easy fight. *Guardian* (July 13), https://www.theguardian.com/music/2016/jul/13/taylor-swift-youtube-music-royalties-battle.

Sorensen AT (2007) Bestseller lists and product variety. *J. Indust. Econom.* 55(4):715–738.

Spahr W (2011) GEMA under fire for royalties dispute with YouTube. *Billboard* (June 24), http://www.billboard.com/biz/articles/news/publishing/1177342/gema-under-fire-for-royalties-dispute-with-youtube.

Sun M, Zhu F (2013) Ad revenue and content commercialization: Evidence from blogs. *Management Sci.* 59(10):2314–2331.

Susarla A, Oh JH, Tan Y (2012) Social networks and the diffusion of user-generated content: Evidence from YouTube. *Inform. Systems Res.* 23(1):23–41.

Takeyama LN (1994) The welfare implications of unauthorized reproduction of intellectual property in the presence of demand network externalities. *J. Indust. Econom.* 42(2):155–166.

Tang Q, Gu B, Whinston AB (2012) Content contribution for revenue sharing and reputation in social media: A dynamic structural model. *J. Management Inform. Systems* 29(2):41–76.

TechCrunch (2007) YouTube launches revenue sharing partners program, but no pre-rolls. (May 4), https://techcrunch.com/2007/05/04/youtube-launches-revenue-sharing-partners-program-but-no-pre-rolls/.

Thomes TP (2013) An economic analysis of online streaming music services. *Inform. Econom. Policy* 25(2):81–91.

Tucker CE, Zhang J (2011) How does popularity information affect choices? A field experiment. *Management Sci.* 57(5):828–842.

Wallop H (2008) CD singles off the shelves at Woolworths. *Telegraph* (May 26), https://www.telegraph.co.uk/news/uknews/2033246/CD-singles-off-the-shelves-at-Woolworths.html.

Warner T (2003) *Pop Music—Technology and Creativity* (Ashgate, Farnham, UK).

Wen W, Forman C, Graham SJ (2013) The impact of intellectual property rights enforcement on open source software project success. *Inform. Systems Res.* 24(4):1131–1146.

Williams HL (2013) Intellectual property rights and innovation: Evidence from the human genome. *J. Political Econom.* 121(1):1–27.

Wlömert N, Papies D (2016) On-demand streaming services and music industry revenues—Insights from Spotify's market entry. *Internat. J. Res. Marketing* 33(2):314–327.

YouTube (2008) Like what you see? Then click-to-buy on YouTube. *YouTube Official Blog* (October 7), https://youtube.googleblog.com/2008/10/like-what-you-see-then-click-to-buy-on.html.

Zentner A (2006) Measuring the effect of file sharing on music purchases. *J. Law Econom.* 49(1):63–90.

Zhang L (2018) Intellectual property strategy and the long tail: Evidence from the recorded music industry. *Management Sci.* 4(1):24–42.

Zhou R, Khemmarat S, Gao L (2010) The impact of YouTube recommendation system on video views. *Proc. 10th ACM SIGCOMM Conf. Internet Measurement* (ACM, New York), 404–410.

Zhou R, Khemmarat S, Gao L, Wan J, Zhang J (2016) How YouTube videos are discovered and its impact on video views. *Multimedia Tools Appl.* 75(10):6035–6058.