Copyright and the Production of Hip-Hop Music

Jeremy M. Watson

*Boston University*

This Draft: June 20, 2017
Copyright and the Production of Hip-Hop Music

Abstract
Whereas the role of patents in cumulative innovation has been well established, little work has examined the impact that copyright policy may have on cumulative innovation in creative content industries. Utilizing U.S. federal court decisions that strengthened the breadth of copyright policy, this paper examines the implications of those decisions on the re-use of original content in the popular music industry, particularly hip-hop music. With a novel, self-collected data-set that tracks re-use through “digital sampling” in hip-hop music, I explore the impact that these federal copyright cases had on the production process of hip-hop music through changes in sampling practices. I find that digital sampling, wherein new musical works are created in part from existing sound recordings, significantly decreased following a 1991 decision that effectively strengthened rights for the original rights holder while restricting re-use. Additionally, I find evidence that this effect on sampling was greater in magnitude for more established artists, and also find evidence that this decision lead to a small decrease in novel works being re-purposed in new songs.

Keywords: Copyright, Intellectual Property, Content Industries

JEL Codes: O34.
1 Introduction

Over the past decades digitization has had a profound effect upon content industries, especially the music industry. The advent of peer-to-peer software (e.g., Napster) and rapid diffusion of content over the internet threatens the excludability of products in such industries, effectively weakening copyright protection. At the same time that digitization threatens content industries through unauthorized file sharing, technological progress ushers in new technologies that lower production costs for new content, as well as creating new, efficient channels for distributing content goods (Waldfogel, 2012). Technological progress can also contribute to the development of new styles and genres within the music industry, such as with the advent of the digital sampler and the emergence of hip-hop music. New production technologies, like the aforementioned digital sampling devices and the more recent digital audio workstations (DAW), decrease the costs of re-use in creating new works that build off past works. However, the intellectual property rights regime may be at odds with the direction afforded by such advances in technology.

Music, like most forms of art, relies and builds upon the existence and ideas expressed in previous works. However, copyright, which protects specific rights associated with the creation of original works, imposes limits upon the ability of artists to re-use prior work in new contexts. Even though U.S. copyright law was established at the federal level with the Copyright Act of 1790, it was only with the Copyright Act of 1976 that a general right was added allowing authors to control derivatives based upon their works. The apparent strengthening of copyright over time has been of growing concern in the music industry, as even absent direct re-use of a sound recording or musical composition, courts may find creators of new works to be infringing the rights of an older vintage. The recent jury decision that Robin Thicke’s single “Blurred Lines” infringed the rights of Marvin Gaye’s “Got To Give It Up” exemplifies this restriction over re-use, despite any direct copying, as does the less recent decision that George Harrison, in writing “My Sweet Lord,” subconsciously plagiarized a work by Ronnie Mack1. While intellectual property rights regimes must weigh the moral rights for ownership as well as the economic incentives faced by potential innovators, concerns over the strength of granted rights are especially relevant when innovation is cumulative. As discussed in Scotchmer (1991), while downstream creators in a cumulative innovation context may bargain with inventors to license upstream rights for re-use, the incentives downstream can often be deficient under strong, restrictive property rights. Even though the impact of patent rights on cumulative innovation has been well explored (see e.g., Green and Scotchmer (1995); Galasso and Schankerman (2015); Williams (2013)), little work has examined copyright’s effect upon

---

innovation when it is cumulative. The lack of research in this domain may perhaps be due to the narrow breadth of copyright, protecting only the expression of an idea, not the idea itself. However, the declining costs of technologies that enable the creative re-use of existing media into remixes and mashups, combined with the new digital distribution channels for such work (e.g., YouTube), has made research in this area all the more important.

This paper, utilizing a novel dataset on a form of explicit re-use in popular music – known as sampling – estimates the effect that recent U.S. copyright court decisions have had on sampling practices in the music industry. By comparing sampling trends before and after precedent-setting court decisions, while shrinking the sample window to exclude time-varying effects, I find that the first important copyright decision lead to a mean drop of between 0.3 and 0.45 songs re-used in each new hip-hop track. While two sampling related court decisions were identified and studied, only the first decision appeared to have a meaningful impact on sampling practices in the industry. Despite the large drop in the magnitude of sampling, there appears to be no change on the propensity to use samples, with the results being driven by the intensive margin. Furthermore, I find that the magnitude of this decline was greater for more prominent artists. Additionally, the results herein suggest that after the 1991 *Grand Upright...* decision, fewer “novel” samples were being used, while re-use of samples from songs that had previously been sampled became more common.

1.1 Digital Sampling

While many genres of contemporary music may incorporate some form of “sampling,” hip-hop music in particular began with a focus on sampling fragments of existing sound recordings to create a new work\(^2\). The roots of modern sampling practices in popular music can be traced back to disc jockeys (DJs) manipulating vinyl records via turntables and crossfaders, using such equipment, DJs could loop over interesting segments of a song, or isolate particular instrumentals, such as the drum “break.” Such techniques extended beyond just replaying parts of the song in a recognizable manner – DJs can manipulate the playback of the record using techniques like “scratching” to create sound effects that are almost unrecognizable as being sourced from the original recording. While such techniques were useful for live performances, the advent of digital sampling devices in the 1980s significantly enhanced producers’ flexibility in recording new, sample-based music. By allowing artists to store, loop over, and sequence samples, the digital sampler helped transform hip-hop into a recorded, commercially-relevant art form.

\(^2\)Sampling is defined in this paper as the re-use of a portion of an existing sound recording or musical composition in a new song or recording.
Musicians that incorporate sampling may view the technique as more than routine appropriation or theft of an original sound recording, but the use of samples runs the risk of infringing at least two rights associated with the sampled work: (1) the copyright of the sound recording, and (2) the copyright of the underlying composition. While early hip-hop musicians up to late-1991 utilized unlicensed samples in a legal gray area (see discussion in McLeod and DiCola (2011)), federal copyright cases since 1991 have set precedent and highlighted the risk of infringement that comes with unlicensed sampling.

A recent theoretical literature has emerged that focuses on models of sampling, (or “remixing”) wherein a downstream sampler bargains with an upstream creator whose original work is repurposed by the downstream agent (DiCola, 2010; Gans, 2015). While such literature is useful in considering the incentive constraints faced by upstream and downstream creators in alternative copyright policy scenarios, this paper builds upon these theoretical papers by providing the first quantitative evidence of how sampling and re-use is affected as copyright policy has effectively grown more restrictive.

1.2 Federal Copyright Decisions on Sampling

1.2.1 Grand Upright v. Warner Brothers Records

While there have been many copyright lawsuits over sampling, with the first notable case regarding the Beastie Boys’ 1986 album License to Ill, many of these cases were settled out of court and hence provided no public precedent to set the contours of copyright policy with respect to sampling. The first case on sampling to be settled via court decision was Grand Upright Music v. Warner Bros. Records, Inc., 780 F. Supp. 182 (S.D.N.Y 1991) (henceforth Grand Upright). This case concerned the rap artist Biz Markie’s use of a sample from a Gilbert O’Sullivan recording on Markie’s album “I Need a Haircut.” Judge Kevin Duffy, in ruling in favor of the plaintiff, Grand Upright Music, granted an injunction against the defendant, with whom Biz Markie was under contract. In his opinion, Duffy stated “Thou shalt not steal,” essentially equivocating the digital sampling of music with theft. While the ruling did not provide an in-depth analysis of copyright law with respect to sampling, the infringement ruling highlighted the risk of sampling in the music industry. Prior to Grand Upright sampling in hip-hop had been likened to lawlessness in the Old West, as by Eothen Alapatt of Stones Throw Records: “it was a kind of Wild, Wild West situation where no one really knew the legalities. Everyone was just doing it.”

3 By strengthening the breadth (or perceived breadth) of copyright in this respect, the ruling may have affected future sampling practices in the recording industry,

---

3As reported in McLeod and DiCola (2011)
a conclusion supported by interviews with industry participants found in McLeod and DiCola (2011). For example, McLeod and DiCola, in quoting entertainment lawyer Whitney Broussard, “For many years after ... they [Warner Bros.] wouldn’t let you release anything that wasn’t licensed. They had a sample committee that would listen to records to see if they could find undisclosed samples. So, they take that pretty seriously.” McLeod and DiCola additionally argue that out-of-court settlements for infringement likely increased after 1991, as a successful defense appeared unlikely after the broad ruling by Judge Duffy. Faced with the outcome of Grand Upright, record producers may either attempt to use unlicensed samples in their music and risk costly infringement claims based on this precedent, or pay lawyers to track down rights holders and bargain over a license to incorporate samples in their work.

1.2.2 Bridgeport Music v. Dimension Films

The second case of interest for this paper, that also set controversial precedent restricting the use of sampling, was Bridgeport Music, Inc. v. Dimension Films, 410 F. 3d 792 (6th Circuit 2005) (henceforth Bridgeport Music). This case, brought by the apparent rights holder for Funkadelic’s “Get Off Your Ass and Jam,” concerned the unauthorized usage of a three-note guitar sound from the Funkadelic recording in N.W.A.’s track “100 Miles and Runnin.” The opinion of the 6th circuit determined that a de minimis defense against infringement did not apply to sound recordings. The court ruled in its opinion, “Get a license or do not sample. We do not see this as stifling creativity in any significant way.” Again, this case anecdotally had a large effect on the advice given by lawyers practicing in the industry, as evidenced by McLeod and DiCola (2011) quoting one music lawyer “I would advise my clients before Bridgeport if they used a little snippet of a recording that was de minimis, 'That’s fine; we dont have to clear it,' ” whereas according to another lawyer post Bridgeport Music “they probably got even more conservative about clearing stuff. Basically, it said that even if you can’t hear a sample of the sound recording, you still have to clear it.”

The remainder of the paper provides a brief review of prior research related to this study (Section 2), develops theory and proposes hypotheses (Section 3), explains the empirical framework, data, and identification strategy (Section 4), presents the empirical results (Section 5), and finally concluding remarks (Section 6).

2 Related Literature

The main contribution of this paper is to uncover how the breadth of copyright policy affects follow-on innovation. Of particular relevance, this paper explores how copyright may contribute to hold-up when innovation is cumulative, and is one of the first papers to examine how copy-
right policy and digitization affects the content of new goods. While economists have been specifically interested in copyright for some time now (see e.g., Landes and Posner (1989)), the economic literature on copyright is relatively sparse compared to the breadth of literature on patents. Much of the recent empirical work on copyright was motivated by the development of file-sharing and peer-to-peer technologies, focusing on the extent to which file-sharing did or did not displace legitimate record sales (Liebowitz, 2006, 2008; Oberholzer-Gee and Strumpf, 2007, 2010; Rob and Waldfogel, 2006). Another major stream of recent empirical work on copyright has focused on the effect of copyright’s duration and implementation on creators’ incentives and the supply of new creative goods (Hui and Png, 2002; Png and Wang, 2006; Li et al., 2013; Giorcelli and Moser, 2014; Liebowitz and Margolis, 2004). From a broader perspective, this paper contributes to our understanding of the effects of intellectual property policy on innovation, of which most has focused on the patent system and academic research (Williams, 2013; Kortum and Lerner, 1999; Sakakibara and Branstetter, 2001; Heller and Eisenberg, 1998), though of particular parallel to this paper is the recent study by Hall and MacGarvie investigating the effect of federal court decisions on the scope of software patents (Hall and MacGarvie, 2010).

This paper, to my knowledge, is the first research to study how copyright policy affects the content of new products, and one of the first to examine copyright in a cumulative context. Closely related work by Nagaraj (2016) has studied digitization and copyright in a cumulative innovation context, utilizing a natural experiment with the Google Books project to estimate the negative impact of copyright on re-use of the Baseball Digest magazine in Wikipedia articles. However, Nagaraj (2016) focuses on upstream works that have fallen into the public domain, whereas this paper examines policy changes that affect the breadth of rights for existing copyrights. Given the unique dataset in this paper, I am additionally able to examine how copyright policy affects changes in commercially released new products.

The focus on sampling in this paper, a production practice in the music industry that was greatly enabled by technological change\(^4\) closely ties this research with the recent work studying digitization and its impact on the creative industries. Research in this area has examined the impact of new, digititally-enabled distribution channels on music industry profits (e.g., Mortimer et al. (2012)), as well as the effect of digitization on the quality of new products in the music industry (e.g., Waldfogel (2012)). An expanding research stream in this area has focused on firms’ efforts to protect their copyright goods under the threat of digital technology that decreases barriers to unlicensed copying (Luo and Mortimer, 2016; Zhang, 2016). Field experiments in stock-photography enforcement and settlement have highlighted the role of communication in protecting IP assets (Luo and Mortimer, 2017), while observational studies

---

\(^4\)beginning with digital sampling devices, now dedicated software
have estimated the positive e-book sales increase that results from enforcement through Digital Millenium Copyright Act (DMCA) takedown procedures (Reimers, 2016).

Due to the cumulative innovation setting of this paper, it builds off the existing theoretical and empirical literature on cumulative innovation and patents, and extends upon this stream by providing empirical evidence of the effect of copyright on follow-on innovation (Green and Scotchmer, 1995; Scotchmer, 1991; Bessen and Maskin, 2009; Hall and Ziedonis, 2001). By examining the rate and direction of innovation after intellectual property rights become more restricted, this paper may also deepen our understanding of the costs and benefits of open-access for innovation in a novel setting (Furman and Stern, 2011; Murray and Stern, 2007).

A major contribution of this paper is to the aforementioned literature focusing on sequential music creation (e.g., sampling, remixes) and the role of the copyright system (McLeod and DiCola, 2011; DiCola, 2010; Menell, 2015; Gans, 2015). This empirical study builds off and complements these prior theoretically-focused examinations, and as the first such empirical research in this vein, provides quantitative estimates of how copyright policy has affected the sampling practices of industry participants.

3 Theory Development and Hypotheses

Licensing of copyrighted works may follow a number of different paths depending upon the rights desired and the context of their desired usage. For some cases, compulsory licensing provisions exist in U.S. Copyright law – such as the right for artists to create “cover versions” of previously recorded musical works. In addition to compulsory licensing schemes, copyright collectives have been created in many countries that represent groups of copyright owners to manage the licensing of selected rights, as well as the collection of royalties. Performance rights organizations (PROs), such as ASCAP and BMI, are a notable example of copyright collectives, that specifically work to license and manage the public performance right (e.g., playing a song in a retail store) for participating copyright holders. However, in most cases, the licensing of copyrights must proceed in an ad hoc, un-standardized manner when prior works under copyright are incorporated into a new derivative work. Negotiations over this right may proceed in either direction, a potential licensee may approach a rights holder to obtain a license to create a derivative, just as a rights holder may identify their unlicensed work in a published derivative work and seek out the infringing artist.

Two separate rights must be licensed in the context of sampling a sound recording, a “mechanical license” for the underlying composition, as well as a “master use” license for the rights associated with the sound recording itself. These two rights are often owned by separate parties, with the sound recording copyright typically controlled by the record label.
or performing artist(s), and the composition copyright typically vesting with the song writers or their publisher. Furthermore, each of these rights may be fragmented, particularly the composition right, as songs are often written by multiple song writers, all of which may be under contract with independent publishers. When these rights are successfully licensed for use as samples, Broussard (1991) identifies five types of agreements that are typically encountered: gratis, buyouts, royalties, co-ownership, and assignment of copyright. Additionally, the form of the license agreement often differs for the right being licensed. While master user licenses typically take the form of "buyouts" (or flat fees), with royalty agreements being less common, royalty agreements and co-ownership agreements are the most common deals for licensing of the composition right. Along with the types of deals negotiated, Broussard (1991) highlights three main factors that influence the price of the negotiated sample license agreement - what is sampled, how the sample is used, and who is using the sample. Typically, in terms of what is sampled, price depends on the popularity of the original artist/song, as well as what part of the song is sampled (e.g., vocal/instrumentals, chorus, melody, etc.) and how recognizable the sampled section is. The negotiated price also depends on how the sample is repurposed in the new work, with higher rates associated with samples that are repeated throughout the song and samples that provide a great deal of the new derivative work's appeal. The final factor determining price, who is using the sample, generally depends on the prominence of the artist using the sample, with very commercially successful artists expected to pay a higher rate. However, with equal importance to "who" uses the sample, is whether the sampling artist sought prior permission to use the sample, with heavy penalty pricing imposed when the original rights holder finds infringing content and begins negotiations post-release.

3.1 Transaction Costs and Hold-up

The above discussion pertains to deals in which licensing terms were agree upon. However, reaching such agreements may not always be possible when upstream property rights are strong. The licensing of rights in this context is complicated by a number of transaction cost hurdles that can lead to hold-up and inefficiency. To begin with, licensing of samples entails significant search costs. If a sound recording is licensed, both sets of rights holders for the master and composition right must be identified and tracked down before negotiations may even begin, and such a process can be far from trivial. Both of these rights may be fragmented due to the often collaborative nature of music production, with the composition right being at particular risk for fragmentation. For example, the recent Billboard chart-topping single "Can't Feel My Face" by The Weeknd has five separate songwriters associated with the underlying composition
While in the best case all songwriters may be under contract with a single major publisher (e.g., BMG, Universal, EMI), the situation is often much bleaker with control split among publishers, or with rights held in part by independent songwriters. Further complications may arise in practice - an artist being sampled may have a contract that requires their permission before their work may be sampled, or the original songwriter and publisher may no longer own the copyright in question.

However, once all of the relevant rights holders are identified, a negotiations process must begin with a number of complications in its own right. Even if rights holders are interested in cooperating with downstream samplers, negotiations may be complicated by the manner that the sample is used in the new work, or the context in which it is used. In worse cases, rights holders may have divergent interests, with a portion open to, or even reaching licensing terms, with some relevant party holding out.

Perhaps of greatest concern, and theoretical source of hold-up, relates to the manner in which licensing negotiations proceed. The transaction cost economics literature (e.g., Williamson (1985)) provides intuition here for how licensing will be affected when upstream rights are strong. Artists must spend a significant amount of effort searching for and experimenting with samples during early production before the necessary samples and potential licensors are identified. A sampling artist thus must expend a considerable amount of effort before negotiations over sample licenses may begin, often creating a new song before negotiations open. Once a sample has been selected by the artist for their new work, their investment in creatively using the sample is highly specific to the relationship with the owners of the sampled copyright. This sunk investment, in combination with ex-post licensing that is typical in practice, opens up a hold-up problem. With incomplete contracts, parties cannot realistically contract over licensing before potential licensees sink their investment in production, which positions downstream licensees at risk of being exploited by upstream rights holders. The first hypothesis follows from this discussion.

*Hypothesis 1. When copyright is restrictive of downstream re-use, the magnitude of re-use will drop.*

### 3.2 Complementary Monopoly

Cournot’s theory of complementary monopoly provides further intuition for how strong upstream rights may affect re-use downstream (Cournot, 1838; Economides and Salop, 1992). Cournot considers the case of two monopolists selling complementary goods (e.g., zinc and...
copper) used in a downstream market. When both of the monopolists act individually to maximize profits, equilibrium prices result that are greater than would be obtained from an integrated monopolist. In addition to the increased prices, total welfare decreases when the upstream producers act independently compared to the integrated case as the independent monopolists do not account for how their own actions affect demand for the other. Such a model is a useful example when considering the licensing of copyrights for producing derivative works (e.g., sample-based music). When upstream rights holders independently price licenses of complementary works, sub-optimal equilibrium may result. In creating derivative works, this situation can arise in two ways in the music licensing context. First, creators of derivative works may have to negotiate with multiple rights holders covering a single work when rights are fragmented. Second, complementarities exist between works when the derivative work is created. That is, after downstream artists invest in creating a derivative work with multiple samples, each upstream rights holder controls a complementary copyright that the downstream producer must license before their derivative work can be commercialized under restrictive re-use. Acting independently, the complementary rights holders price their licenses above the level obtained in an optimal equilibrium. In response to sub-optimally high prices, it is expected that creators of derivative works will decrease the number of samples used in order to alleviate costs, but this effect will be obtained through a decrease in number of samples used per song, rather than a decrease in the number of works containing samples due to the established norms of re-use. Thus leading to the following hypothesis.

*Hypothesis 2. When copyright restricts downstream re-use, fewer samples will be used per derivative work, compared to a regime of more permissive rights*

### 3.2.1 Heterogeneity by Prominence

Given the above discussion, we may also expect the proposed effect on sampling to have differential impact across artist types. For example, artists and labels may differ in their willingness to infringe, or risk infringing, the copyrights of others through unlicensed sampling. Furthermore, artists may have varying ability to clear transaction cost hurdles and negotiate licenses. To focus in on this question, heterogeneity in the sampling effect according to artist prominence is explored. However, the theoretical expectation in this situation remains particularly unclear. On the one hand, high profile may have the resources necessary to clear licenses when rights are restrictive. Such resources available to prominent artists could include higher production budgets, as well as the availability of administrative and legal support for licensing samples. On the other hand, prominent artists have much greater exposure, and unlicensed sampling may be easily identified by upstream rights holders. An analogous effect has been observed in
the patent space, with non-practicing entities (NPE) litigation preferentially targeting cash-
rich firms (Cohen et al., 2014). Because of this, major labels with which prominent artists are
associated may be highly restrictive over artists’ use of samples, and may monitor their releases
closely to ensure they are not at risk of costly litigation. With these predictions in mind, it
is uncertain which effect will dominate, and thus whether prominent artists will be more or
less affected by restrictive sampling rights than less prominent artists. The exploration of this
question thus depends on the empirical results in the following sections.

4 Empirical Framework and Identification

4.1 Data and Measures

Data for this paper was collected from two sources, WhoSampled.com and Billboard. Data on
music sampling was self-collected from the website WhoSampled.com. WhoSampled, which bills
itself as “the world’s largest community for fans of sampled music, cover songs and remixes,”
is a community-driven website where contributors upload information about samples used in
songs. The database has information on 365,330 unique songs and 207,632 samples, provided
by 14,230 contributors at the time of this draft. While the WhoSampled database also has
information on cover songs and remixes, for the purposes of this study only sampling data was
used. For each sample driven song in the database, WhoSampled provides a number of fields
important to building this dataset: the artist, title, label, and date of release of the sampling
(“follow-on”) song, plus the artist, title, label, and date of the sampled song. Additionally, the
WhoSampled database contains information on the genre of both the sampled and sampling
song, as well as where in the song the sample is used, and where the sample is taken from in the
original track. While the “importance” or “prominence” of the sample in the sampling song
would be a useful measure for this study, the WhoSampled data unfortunately does not provide
any such measures. Figure 1 depicts the typical track level data available on WhoSampled for
a sample-using song. In addition to the high level details shown in Figure 1, the community
provides additional detailed information for each specific sample used in a song, as shown in
Figure 2.

The second source of data used for constructing this study’s dataset comes from Bill-
board.com. Billboard provides weekly music industry top-charts for singles, and now ranks
singles based upon digital sales, physical sales, radio play, and online streaming. For this
paper, data from Billboard’s “Hot R&B/Hip-Hop” charts were collected since Billboard be-
gan tracking this genre in the beginning of 1985. Each weekly chart contains a ranked list of
Hip-Hop/R&B singles, including the title of the ranked song and the song’s artist(s).
To construct the dataset used herein, the list of hip-hop/R&B artists appearing on the Billboard charts since 1985 is merged with the artists field of the WhoSampled data. The dataset then includes all sampling tracks found in the database since 1985. The main response variable, \( \text{samples} \) is constructed by calculating the count of samples used in the sampling song. An additional response variable, \( \text{new samples} \), is constructed by calculating whether the sampling song \( i \) has any novel samples - samples of songs for which song \( i \) is the first sampling track - and equals 1 if song \( i \) contains 1 or more novel samples, 0 otherwise. Two treatment variables were created, \( \text{post-Grand Upright} \) which is a binary indicator variable that equals 1 if song \( i \) was released after \( \text{Grand Upright} \), and \( \text{post-Bridgeport Music} \) that analogously equals 1 if song \( i \) was released after \( \text{Bridgeport Music} \). Summary statistics for this data are shown in Table 1.

4.2 Estimation

To investigate the effect that a court decision has on the sampling practices in the music industry, I first focus on the raw count of samples used in a new song. In an ideal experiment to study this question, the econometrician would expose a randomly selected treatment group of sample-using musicians to more (or less) restrictive copyright policy for re-use of existing work. The econometrician could then compare the output of this treatment group to a control group that was not exposed to a change in policy. With this ideal setup, the econometrician could then simply estimate the change in the rate of sampling that is associated with stronger/weaker copyrights over re-use.

Unlike the ideal experiment, the identification strategy in this paper must deal with the fact that there is little heterogeneity in the treatment condition. As during the period of the first court case, \( \text{Grand Upright} \), sampling was primarily found in hip-hop music, and the vast majority of this music was being produced in the U.S. While assumptions about either pre-treatment sampling rates, or treatment heterogeneity based upon jurisdiction of the court could be used to help identify the parameter of interest (e.g., using a differences-in-differences design), this paper takes a simple, transparent approach rather than relying on spurious assumptions given the context. To estimate the effect of copyright court decisions on the rate of sampling, the rate of sampling is compared for songs released before and after the court decision. This pre/post strategy however raises concerns about potential confounding effects from time-trends, as any time-trend will be absorbed into our parameter estimate and bias the results. To deal with this, a shrinking time window is used around the court decision “treatment,” to rule out confounding time trends. While this identification depends upon a maintained assumption regarding time trends within the window, similar strategies have been used in the past with
success (e.g., Zhang and Zhu (2011)), and this strategy is suitable given the lack of appropriate control group. With a linear model, the following regression is used:

\[ s_i = X_i'\beta + \gamma 1[i \text{ is post-court}] + \epsilon_i \] (1)

Where \( s_i \) is a count of samples in song \( i \), \( 1[i \text{ is post-court}] \) is a dummy indicator variable equal to 1 if the song is released after the court decision, and \( X_i \) is a vector of control variables. The available control variables are categorical variables for the main artist associated with song \( i \) and the label of song \( i \). Since sampling was predominantly used in hip-hop music during the first court decision, the entire sample is restricted to the hip-hop music industry to keep the analysis consistent. The same identification strategy and framework shown in Equation 1 is used when investigating the impact of federal court decisions on the incidence of new samples being used in popular music.

5 Results

The results of this analysis are presented in three parts. The first part of this section presents the estimates of the main treatment, the effects of Grand Upright and Bridgeport Music on the rate of sampling in hip-hop music. The analysis proceeds by presenting evidence of how Grand Upright decreased the number of samples used per song, without any significant drop in the probability of a song using samples. I conclude by examining how this main effect differs by artist type, and how the decisions affected the diversity of works being sampled. The majority of the analysis in this section focuses on the Grand Upright case, as will be discussed below, due to the lack of effect that can be attributed to the later court decision Bridgeport Music.

5.1 The Rate of Sampling in Hip-Hop Songs

Figure 3 shows the mean samples-per-song time trend since 1986, shown as the estimates of a fitted Poisson model. From Figure 3 we see that the average number of samples (per-song) peaks between 1989 and 1990, before beginning a sharp decline in 1991, a trend that continues until leveling off in the late 1990s - early 2000s. Such evidence lends support to the claim of Grand Upright’s effect on sampling practices, consistent with Hypothesis 1, but of note is the seeming lack of effect seen in the mid 2000s, where we would expect Bridgeport Music’s effect to appear. Some insightful comparisons can be made between the sampling trend shown in Figure 3 and the past trends in cover songs (Figure 4) and the rate of self-sampling (Figure 5). While such comparisons are imperfect control groups, they give one an idea of whether there were other general trends in re-use over this time period. Figure 4 shows the rate of cover songs per year
as a total proportion of all songs released. Because cover songs are covered by a statutory licensing regime, the incentives to create cover songs should have been completely unaffected by the *Grand Upright* decision, and this does appear to be the case, as the rate of cover songs produced actually increased after 1991, before reaching a peak in the mid 2000s. Second, we can also examine trends in self-samples, that is, samples in which an artist repurposed part of their previous works. As shown in Figure 5, the decline in self-samples post-1991 was much lower than the decline in general sampling. It is of potential concern that any change in self-sampling is observed, but it must be re-iterated that ownership of musical copyrights are often fragmented. One could expect to see some effect of the court-imposed restrictions on self-samples, as just because an artist is associated with a work does not indicate that they control all of the rights needed to license it for re-use.

Table 2 presents the estimated change in samples per hip-hop song after *Grand Upright*. Using a 3 year window before and after the *Grand Upright* decision, the estimated effect on sampling ranges between 0.461 and 0.419 fewer samples per song after the decision, and such effects are significant at the 1% level. While columns (2) and (3) of Table 2 control for potential unobserved heterogeneity with label and artist effects, the longer time window on either side of the *Grand Upright* decision may allow for underlying time-trends in sampling practices to bias the parameter estimates. To lessen concerns about confounding time trends, the time window around *Grand Upright* shrinks to ±2 years in columns (4)-(6) of Table 2. The inference remains robust and largely the same as before, but the estimated negative effect of *Grand Upright* on sampling shrinks slightly in magnitude, to between 0.276 fewer and 0.318 fewer samples per song, the latter estimate controlling for artist and label effects. These consistent results, in total, lend support to Hypothesis 1.

The analysis of *Bridgeport Music*’s effect on the magnitude of sampling in hip-hop is presented in Table 3. In columns (1)-(3), using a time window of ±3 years around *Bridgeport Music* yields a small, negative point estimate for the effect of *Bridgeport Music* on sampling, but with a lack of statistical significance across all models at typical levels. Once the time window shrinks to ±2 years in Columns (4)-(6) of statistical Table 3, the effect of *Bridgeport Music* effectively disappears, with point estimates being further attenuated towards zero, and all statistical significance dropping from the model. Overall, due to the small mean effect size across models, and the lack of statistical significance across models, it appears that if *Bridgeport Music* did in fact have an effect upon the rate of sampling in the industry, such effect is smaller than what can be detected given noise in our sample. Poisson models were also used to estimate the effects in this section, but as this did not change inference the results were omitted.
5.2 Impact on the Probability of Sampling

To further disentangle whether *Grand Upright*’s effect manifested as a change in the proportion of songs that contained samples, or a decrease in the magnitude of sampling across all songs, I construct a dependent variable equal to one if song $i$ contains any samples, 0 otherwise. Running a logit regression using this outcome measure, as shown in Table 4, provides evidence that the court decision did not have an outright effect on the propensity of songs to contain samples, as we cannot reject the null at typical significance levels across the specifications used. It thus appears that the reduced sampling after *Grand Upright* was due to a reduction in the number of samples per track, not the proportion of songs containing samples after the decision, a result consistent with Hypothesis 2. Such results also help to rule out an alternative interpretation that sampling was a fad that peaked in the 1990s then faded from popularity, because one does not see any decrease in the propensity to sample. Instead, one observes the decrease in sampling coming from the intensive margin, an interpretation that is consistent with royalty stacking due to more restrictive copyright.

5.3 Change in Sampling Heterogeneity by Artists

As we cannot reject the null of no effect for most models estimating the impact of *Bridgeport Music*, as shown in the previous section, the rest of the empirical analysis in this paper focuses upon the *Grand Upright* decision and its effect upon sampling practices in the industry. Table 5 presents the results of examining how the *Grand Upright* decision affected different types of artists. By defining musicians as prominent if they had multiple appearances on the Billboard platinum album charts, we see that the change in sampling practices post-*Grand Upright* was greater in magnitude for these established artists, as the interaction between the post-*Grand Upright* indicator and the indicator for being a multiplatinum artist is significant across all specifications. While the established artists not only were more affected by the *Grand Upright* decision, these artists also had a higher baseline rate of sampling per song pre-*Grand Upright* than less prominent artists.

Such results clarify the theoretical discussion in Section 3. Whatever resource advantage these prominent artists may have in clearing samples appears to be dominated by the increased scrutiny and litigation risk they face when using samples that may be unlicensed.

5.3.1 Impact on diversity of samples

The remainder of the empirical analysis in this paper focuses on how *Grand Upright* affected the diversity of what is being sampled. If *Grand Upright* made copyright more restrictive for artists wishing to re-use existing sound recordings, it may have pushed artists and labels to
increasingly re-use work from rights-holders that either (1) did not assert their rights against derivative works, and/or (2) were apt to license their rights to hip-hop producers. Table 6 present results with a new outcome variable that is equal to 1 if the observed song has a sample of a song that was never previously sampled, and 0 otherwise. For concision all results in this section reflect linear probability models estimates - logit models were also estimated but the main inference did not differ.

Table 6 shows the effect of Grand Upright on the incidence of novel samples in hip-hop songs. As seen from the regression results, there appears to be a small negative effect of Grand Upright on novel samples being used to produce new music, with songs post-Grand Upright on average being between 5 and 10% less likely to have novel samples after the court decision. Such results lend support to the interpretation that the downstream restrictions from Grand Upright are leading artists to exploit samples from a less diverse pool of artists and songs than would have been obtained under a less-restrictive regime.

6 Conclusion

The results documented here provide an examination of the impact of strengthened copyright breadth on re-use in the music industry. While the results of the analysis presented here are mixed in interpretation, and require some small maintained assumptions, this work provides evidence of how copyright policy may actually affect the content of new creative products. Future work in this area could further explore this question, as well as focus on the other side of the sample-sampling dyad - the upstream rights holders.

First and foremost, the empirical results indicate that there was a moderate decrease in sampling following the Grand Upright decision. Depending on the model used and time window around the event, the court decision is estimated to have caused sampling to decrease by approximately 0.3-0.4 samples per song on average, and the magnitude of this result is robust across models. An interesting result shown herein lies with the lack of a significant effect coming from the Bridgeport Music decision. While Bridgeport Music had a more dramatic effect on copyright policy itself, as it effectively changed policy within its jurisdiction (see McLeod and DiCola (2011)), it appears that the earlier Grand Upright decision had a more significant impact upon sampling and licensing practices within the industry, as this case may have forced labels to confront the risks associated with unlicensed sampling, while clearing the way for infringement claims. Thus, while this paper fails to find any economically significant or statistically significant effect of Bridgeport Music on sampling in the industry, there does appear to be robust evidence that Grand Upright decreased the rate of sampling and re-use. Strikingly, the evidence shows that the court decision mainly affected derivatives and re-use through an effect on the intensive
margin, with the magnitude of samples used per track dropping, with no apparent change in the probability of a new song containing samples. This effect highlights the royalty stacking problem in licensing for derivative works - in response to multiple upstream rights holders (complementary monopolies), artists decrease the number of samples used in order to offset the suboptimally-high royalty rates charged. Thus, I find not only that the magnitude of sampling decreases when re-use is restricted, but also find evidence of the mechanism at work.

Empirical evidence also appears to indicate that Grand Upright had a differential impact on more/less prominent artists, with the regression analysis providing evidence that more prominent artists were affected to a greater extent by the court decision than other, less prominent artists. Such a result is particularly interesting due to the theoretical ambiguity of how stricter rights over re-use may effect artists. On the one hand, it may be that more prominent artists signed to major labels have the resources to license samples that more independent, less prominent artists do not have. On the other hand, a separate effect can be theorized wherein less-exposed artists (“underground artists”) are able to conduct unlicensed sampling without threats of infringement suits, and are thus able to continue sampling at a higher rate than artists with greater exposure. However, from the empirical results shown in this paper, I see no evidence of the former effect, instead finding that more prominent artists were affected moreso by the decision. Evidence of an effect on the types of work being sampled are presented, with results supporting the claim that novel samples were less likely to be used post Grand Upright.

The study of copyright strength in this context is a complicated issue, necessitating concerns over both moral rights as well as economic incentives. This paper does not attempt to provide a final answer on the matter or to calculate a welfare estimate. Instead, this study seeks to chip-off and answer one question - how did more restrictive rights over re-use affect the content of follow-on work? By answering this question, this paper may contribute to the discussion forming around copyright and the creation of derivative works, a spreading phenomenon enabled by digitization.
References


Appendix A: Figures and Tables

Figure 1: WhoSampled example

This screenshot depicts the community-provided sampling data from WhoSampled.com for the Beastie Boys song “Car Thief”
Along with the high-level sampling data shown above, the community at WhoSampled also provides detailed information for each sample added to the database, shown here for the Beastie Boy’s sample of Funk Factory’s *Rien Ne Va Plus*.

### Table 1: Summary Statistics

<table>
<thead>
<tr>
<th>Statistic</th>
<th>N</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>year</td>
<td>26,892</td>
<td>2,001.657</td>
<td>8.253</td>
<td>1,985</td>
<td>2,015</td>
</tr>
<tr>
<td>samples</td>
<td>26,892</td>
<td>1.223</td>
<td>1.619</td>
<td>0</td>
<td>36</td>
</tr>
<tr>
<td>postGU</td>
<td>26,892</td>
<td>0.860</td>
<td>0.346</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>postBM</td>
<td>26,892</td>
<td>0.386</td>
<td>0.487</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>new samples</td>
<td>19,425</td>
<td>0.349</td>
<td>0.477</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>
This figure illustrates time trend of samples used per song (shown as fitted Poisson-model marginal effects). The time of the *Grand Upright* decision (December 1991) is shown in red.
This figure illustrates time trend of cover songs per year, as a proportion of total songs released per year.
This figure illustrates time trend of self-samples used per song (shown as fitted Poisson-model marginal effects). The time of the *Grand Upright* (December 1991) decision is shown in red.
**Table 2**: *Grand Upright* effect on sampling. This table displays coefficients from OLS regressions in which the dependent variable is a count of samples used in song $i$. The main variable of interest, post-GrandUpright =1 for songs released after the *Grand Upright*... court decision that restricted re-use. Columns (1), (2), and (3) utilize a ±3 year window around the event to minimize confounding time trends. Column (2) controls for record label effects, while Column (3) controls for both artist and label. Columns (4), (5), and (6) shrink the time window further to ±2 years, with Column (6) fully controlling for label and artist effects.

<table>
<thead>
<tr>
<th></th>
<th>±3 Year Window</th>
<th>±2 Year Window</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>post-GrandUpright</td>
<td>-0.461***</td>
<td>-0.443***</td>
</tr>
<tr>
<td></td>
<td>(0.060)</td>
<td>(0.066)</td>
</tr>
<tr>
<td>Intercept</td>
<td>2.096***</td>
<td>1.635***</td>
</tr>
<tr>
<td></td>
<td>(0.048)</td>
<td>(0.489)</td>
</tr>
</tbody>
</table>

| Artist Effects  | No             | No             | Yes            | No             | No             | Yes            |
| Label Effects   | No             | Yes            | Yes            | No             | Yes            | Yes            |
| Observations    | 5,706          | 5,706          | 5,706          | 4,119          | 4,119          | 4,119          |
| Adjusted $R^2$  | 0.010          | 0.118          | 0.202          | 0.004          | 0.142          | 0.198          |
| F Stat          | 58.990***      | 2.652***       | 2.344***       | 18.445***      | 2.888***       | 2.185***       |

*Note:* $^\ast p<0.1; ^{**}p<0.05; ^{***}p<0.01$
### Dependent variable:

**Number of Samples**

<table>
<thead>
<tr>
<th></th>
<th>±3 Year Window</th>
<th>±2 Year Window</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>post-BridgeportMusic</td>
<td>-0.020</td>
<td>-0.067*</td>
</tr>
<tr>
<td></td>
<td>(0.032)</td>
<td>(0.038)</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.967****</td>
<td>1.528****</td>
</tr>
<tr>
<td></td>
<td>(0.023)</td>
<td>(0.100)</td>
</tr>
<tr>
<td>Artist Effects</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Label Effects</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>5,772</td>
<td>5,772</td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>-0.0001</td>
<td>0.102</td>
</tr>
<tr>
<td>F Statistic</td>
<td>0.399</td>
<td>1.690***</td>
</tr>
</tbody>
</table>

**Note:** *p<0.1; **p<0.05; ***p<0.01

Table 3: *Bridgeport Music* effect on sampling. This table displays coefficients from OLS regressions in which the dependent variable is a count of samples used in song $i$. The main variable of interest, post-*BridgeportMusic* =1 for songs released after the *Bridgeport Music* court decision in 2005 that restricted re-use. Columns (1), (2), and (3) utilize a ±3 year window around the event to minimize confounding time trends. Column (2) controls for record label effects, while Column (3) controls for both artist and label. Columns (4), (5), and (6) shrink the time window further to ±2 years, with Column (6) fully controlling for label and artist effects.

### Dependent variable:

**ContainsSamples (1/0)**

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>post-GrandUpright</td>
<td>0.127*</td>
<td>0.040</td>
<td>0.113</td>
</tr>
<tr>
<td></td>
<td>(0.074)</td>
<td>(0.093)</td>
<td>(0.135)</td>
</tr>
<tr>
<td>Intercept</td>
<td>1.006****</td>
<td>0.893</td>
<td>19.453</td>
</tr>
<tr>
<td></td>
<td>(0.058)</td>
<td>(0.838)</td>
<td>(10,753.990)</td>
</tr>
<tr>
<td>Artist Effects</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Label Effects</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Observations</td>
<td>4,119</td>
<td>4,119</td>
<td>4,119</td>
</tr>
<tr>
<td>Log Likelihood</td>
<td>-2,325.973</td>
<td>-1,709.385</td>
<td>-1,196.570</td>
</tr>
</tbody>
</table>

**Note:** *p<0.1; **p<0.05; ***p<0.01

Table 4: Propensity to Sample - Logit on Sampling Dummy: 2 Year Window. This display displays coefficients from a logit regression in which the dependent variable binary (1/0) variable =1 when song $i$ contains samples, 0 otherwise. The main variable of interest, post-*GrandUpright* =1 for songs released after the *Grand Upright* court decision that restricted re-use. Column (1) includes the main variable of interest, Column (2) adds label effects, and Column (3) controls for artist effects.
### Table 5: Multiplatinum Artists: 2 Year Window

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>Number of Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>multiplatinum:post GrandUpright</td>
<td>$-0.855^{***}$</td>
</tr>
<tr>
<td></td>
<td>(0.177)</td>
</tr>
<tr>
<td>multiplatinum</td>
<td>$1.824^{***}$</td>
</tr>
<tr>
<td></td>
<td>(0.145)</td>
</tr>
<tr>
<td>post GrandUpright</td>
<td>$-0.204^{***}$</td>
</tr>
<tr>
<td></td>
<td>(0.076)</td>
</tr>
<tr>
<td>Constant</td>
<td>$1.764^{***}$</td>
</tr>
<tr>
<td></td>
<td>(0.059)</td>
</tr>
</tbody>
</table>

| Artist Effects      | No | Yes | Yes |
| Label Effects       | No | No  | Yes |
| Observations        | 4,119 | 4,119 | 4,119 |
| Adjusted R$^2$      | 0.061 | 0.218 | 0.200 |
| F Statistic         | 89.645$^{***}$ | 3.030$^{***}$ | 2.201$^{***}$ |

*Note:* $^{*}p<0.1; ^{**}p<0.05; ^{***}p<0.01$

This table displays coefficients from OLS regressions in which the dependent variable is a count of samples used in song $i$. These regressions include the previous variable of interest, post-GrandUpright = 1 for songs released after the Grand Upright... court decision that restricted re-use. Additionally, an interaction term is added using the variable multiplatinum, which = 1 when artist $j$ has had multiple platinum album certifications according to the RIAA. A ±2 year time window around Grand Upright... is used for all the models in this table. Columns (1) introduces the multiplatinum variable and the interaction term, Column (2) controls for label effects, and Column (3) controls for both label and artist.
Table 6: Novel Sampling, ±3 Year Window. This table displays coefficients from OLS regressions in which the dependent variable is a binary variable, *Contains Novel Sample(s)*, which =1 when song *i* contains samples that have not been previously used, 0 otherwise. These regressions include the main variable of interest, post-*GrandUpright*=1 for songs released after the *Grand Upright...* court decision that restricted re-use. Columns (1) includes no controls, Column (2) controls for label effects, and Column (3) controls for both label and artist.

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>post-<em>GrandUpright</em></td>
<td>-0.040***</td>
<td>-0.054***</td>
<td>-0.102***</td>
</tr>
<tr>
<td></td>
<td>(0.015)</td>
<td>(0.018)</td>
<td>(0.023)</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.393***</td>
<td>0.321**</td>
<td>-0.061</td>
</tr>
<tr>
<td></td>
<td>(0.012)</td>
<td>(0.125)</td>
<td>(0.702)</td>
</tr>
<tr>
<td>Artist Effects</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Label Effects</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>4,195</td>
<td>4,195</td>
<td>4,195</td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.001</td>
<td>0.067</td>
<td>0.131</td>
</tr>
<tr>
<td>F Statistic</td>
<td>6.674***</td>
<td>1.835***</td>
<td>1.774***</td>
</tr>
</tbody>
</table>

*Note:*  *p<0.1; **p<0.05; ***p<0.01*