Digital Sampling, Copyright Assertion and Creative Reuse

Timothy Simcoe

Boston University, Questrom School of Business and NBER

Jeremy Watson

University of Minnesota, Carlson School of Management

April 20, 2019

Digital Sampling, Copyright Assertion and Creative Reuse

PRELIMINARY

Abstract

Digital sampling involves the recombination of short excerpts from previously recorded music as part of a new work. Although the status of sampling under copyright law has been highly contested, in the United States, most artists have sought permissions to sample prior work at least since the 1991 *Grand Upright Music, Ltd v. Warner Bros. Records Inc.* decision. This paper examines the enforcement strategies of two prominent copyright assertion entities that emerged following *Grand Upright*, and the changing patterns of reuse that emerged during the transition from free to rights protected sampling. We explore similarities between patent and copyright assertion strategies, and show that enforceability was associated with a decline in recombination, but also an greater dispersion in the distribution of sampled work, leading to a "long tail" of sampled music.

1 Introduction

The enforcement of intellectual property (IP) rights is an essential mechanism in IP-based markets for ideas and expressions. Intellectual property regimes provide IP owners the standing to litigate their property rights in order to enjoin infringing uses. The statutes that allow creators and inventors of IP to exclude competitors from their markets have also given rise to non-practicing entities (NPEs), popularly called patent and copyright trolls. NPEs are distinguished by their role as intermediaries in IP markets. Rather than acting as inventors or artists, NPEs target practicing entities that have committed to an infringing act in a market. Much research on patent assertion entities (PAEs) argues that NPE activity is societally harmful, raising the cost to innovate through legal fees and settlement costs while not increasing incentives to innovate (Bessen and Meurer, 2014). However, opponents of this view argue that PAEs confer efficiency gains to markets for ideas by facilitating patent exchange and the licensing of innovations (McDonough, 2006). The vast majority of the NPE literature has focused on patent trolls (Lanjouw and Schankerman, 2004; Bessen et al., 2011; Cohen et al., 2016), but these studies are limited by their inability to examine the *ex post* demand for technology rights asserted by trolls. With data on music sampling in this paper, we are able to directly measure the use of litigated IPR in a manner that is not possible with patents.

Copyright trolling activity has recently increased, incentivized by copyright's statutory damage provisions and enabled by digital technologies that decrease the costs of reproducing, distributing, and subsequently identifying infringing content. While copyright's statutory damages are intended to avoid burdening rightsholders with complex or impossible calculations of lost profits from infringement, these same provisions incentivize litigation by NPEs who may have little actual damages from infringement due to their non-practicing nature. Similar concerns regarding NPEs' impact on economic efficiency, initially raised by the PAE literature, have been applied to copyright trolls (DeBriyn, 2012; Balganesh, 2013; Sag, 2015). This paper examines the impact of copyright trolls in the music industry in order to study litigation's effect on patterns of re-use for copyrighted material.

We contribute to the NPE debate by providing the first empirical study of copyright trolls' effects on re-use of prior works through a phenomena of "digital sampling." Digital sampling, henceforth sampling, is the practice of taking a section of a past sound recording for re-use in a new recording. The commercialization of digital sampling devices in the 1980s enabled the widespread use of sampling in hip hop music, in which early samplers held uncertain beliefs

over the copyright infringement or fair use implications of their use. Copyright litigation, in part by the firms studied in this paper, helped trace the contours of property rights in this context, and established that most unlicensed sampling in the U.S. will be regarded as copyright infringement. Our unique ability to observe original source material and subsequent products that incorporate a given source, or sample, allows us to trace trolling's effect on the commercial exploitation of creative material. These results thus help to distinguish between NPEs' potential roles as market makers in IP markets versus their acquisition and litigation strategies inducing an overall negative effect on the exploitation of acquired works and technologies.

2 Copyright and Patent Assertion Entities

A guiding concern in the patent assertion literature is the extent to which the private rewards to asserting a given intellectual property right, especially a patent, diverge from the social value contributed by the invention covered by said patent (Scott-Morton and Shapiro, 2016). To what extent do these concerns extend to copyright systems? The answer to this question is not immediately clear given that the scope of a copyright is narrowly defined compared to patent rights, that copyright regimes typically provide limitations and exceptions to the exclusion right (i.e., assertion rights) of copyright owners, and that the duration of a copyright is much longer than that of a patent.

Digital technology innovations over the past several decades have led to an expansion of activity by patent assertion entities. The internet has decreased search costs for connecting patent inventors with patent intermediaries, while the rise of IP-intensive industries that rely upon information and communication technologies has provided PAEs with ample targets for litigation (Hagiu and Yoffie, 2013). These same technologies have also led to a parallel increase in assertions by copyright owners. Copyright owners and content creators, through use of the internet, may now easily find an aggregator or assertion entity that will enforce their copyrights. The decrease in replication costs due to digitization has increased the incidence of re-use, providing targets for assertion entities. Many uses may fall under fair use, but it is unclear whether rightsholders consider the context of use before asserting their property rights. Search engines, tracking of peer-to-peer networks, and content recognition software may now also be utilized by asserting entities to identify new digital content that infringes their copyrights.

What is a Troll?

Several features of the U.S Copyright Act of 1976 encourage and enable the litigation strategy of firms specializing in copyright assertion. Under 17 U.S. Code § 501, the owner of a copyright, or an exclusive licensee to any of the six rights specified under 17 U.S. Code § 106 may bring action against infringement of any particular controlled right. The owner of a copyright has the exclusive rights to (1) reproduce the copyrighted work, (2) prepare derivative works, (3) distribute copies of the work, (4) to perform the work in public, (5) to display the work in public, and (6) to perform the work in public via digital audio transmission. The exclusive licensee of any of these six rights is considered an owner with respect to infringement of the exclusively licensed right. However, these rights cannot be infinitely divided with e.g., an exclusive license to bring lawsuits enforcing a copyright. Secondly, plaintiffs may seek statutory damages from infringement, and need not prove actual lost profits due to the litigated use.

Lemley and Melamed (2013) separate patent trolls into three distinct business models, "lottery-ticket" trolls, "bottom-feeders," and "patent aggregators." Lottery-ticket trolls own a potentially high-quality patent that covers an important technological area, and attempt to assert this patent in court against industry incumbents towards an end goal of large jury awards for damages. "Bottom-feeders," on the other hand, often hold many low quality patents that are at high risk of invalidation in court. These plaintiffs file "nuisance" lawsuits against a large number of defendants, aimed at recovering small settlements from many targeted firms under the implicit threat of high costs of defensive litigation. Even if the defendant is very likely to win these defendants, it may be cheaper to settle than pay costs of patent litigation. Patent aggregators, the third type of troll, amass a large numbers of patents and license their portfolio to firms under a threat of litigation for those firms that refuse.

Many copyright infringement lawsuits are characterized by creators suing other creators in an isolated manner, arguing that a particular work of the defending artist unfairly appropriated or plagiarized the plaintiff's creations. Regardless of the validity of the plaintiff's claims, these lawsuits parallel the lottery-ticket patent lawsuits, as they often feature a low-profile artist suing a successful artist who is accused of not giving due credit or royalties to the plaintiff.

Litigation by copyright collective rights organizations (CROs) may resemble the activity of patent aggregators in several ways. A common form of CRO is the performance rights organization (PRO), a collective that manages licenses and collects royalties for the performance rights of its members' copyrights. The American Society of Composers, Authors, and Publish-

ers (ASCAP), one of three PROs in the United States, represents over half a million members and more than ten million associated copyrights. PROs alleviate the transaction costs of connecting potential licensees with rightsholders, while also solving the multiple marginalization problem of negotiating separately with each licensor.[1] PROs have been criticized for their aggressive copyright enforcement and litigation against potential licensees, often small businesses, against which PROs leverage their large portfolio of potentially infringed copyrights.[2]

Accusations of "sample trolling" have arisen in response to the aggregation and enforcement of music copyrights against sampling in hip hop and rap music.[3] Bridgeport Music and Tuff City Records, studied in this paper, aggregate and acquire the copyrights to old musical recordings and compositions that have been heavily used as samples in hip hop music. Many of the lawsuits by Bridgeport and Tuff target infringing uses of music released a decade or more prior to litigation. This practice of aggregating copyrights for assertion in mass lawsuits against previously released works has led to accusations of opportunistic behavior, but these "trolling" labels have contended that they act as intermediaries in copyright markets.

Content industry trade organizations, like the Recording Industry Association of America (RIAA) and Motion Picture Association of America (MPAA), organized mass copyright enforcement and litigation against consumer piracy after the advent of peer-to-peer filesharing networks. Rather than targeting businesses as PAEs do, these mass lawsuits have often targeted individual consumers accused of piracy and associated copyright infringement on the internet. Although it has since ramped down the intensity of its litigation campaign, the RIAA asserted copyrights against more than 30,000 individuals from 2003 to 2008 via John Doe lawsuits.[4] The RIAA and its members commonly served targeted consumers with pre-litigation letters, demanding several thousand dollars per work infringed under the threat of statutory damages if the assertion proceeded to litigation. Despite this, it is unclear that this litigation campaign was ever intended as a source of profit for the RIAA's members, or instead an effort to curb rampant online piracy that had been blamed for slumping music industry revenues.

Copyright owners and aggregators have taken interest in the nuisance value, and potential

---

[1] e.g., A restaurant that desires to play background music has little use for licensing just one song to play, but must license a variety of songs with distinct copyright owners.

[2] https://hudsonvalleyone.com/2012/04/09/popular-open-mic-pay-2k-licensing-shut/ (Accessed 4/25/2018)

[3] http://www.slate.com/articles/arts/culturebox/2006/11/jayz_versus_the_sample_troll.html (Accessed 5/3/2018)

[4] https://www.eff.org/wp/riaa-v-people-five-years-later (Accessed 4/25/2018)

profit, of mass assertion campaigns designed to extract settlement payments from targets under the threat of costly IP litigation. Copyright aggregators, such as Righthaven LLC, engage in mass lawsuits filed against hundreds of defendants, threatening purported infringers with statutory damages unless a settlement fee is paid (Balganesh, 2013). The nuisance value and leverage that these assertions hold may be enhanced when defendants are threatened with public exposure for violating copyrights of socially stigmatized content – particularly pornography (Sag, 2015).

A distinguishing feature of patent assertions compared to copyright assertions is the relative risk of unintentional infringement. A firm may develop a new technology, unaware that its invention infringes upon the patent rights of past inventors. This type of "innocent" or unintentional infringement appears much more limited for copyrightable works, though there is certainly a spectrum of infringement. In some cases, despite being cognizant of their use of a work, a creator in a copyright regime with fair use statutes may believe that their use is fair and non-infringing upon the rightsholder. On the other end, the now commonplace peer-to-peer piracy of works is almost certainly witting infringement given the press received by the aforementioned litigation campaigns of the music and movie industries. However, the spread of the internet and digital technologies that diminish reproduction costs has also led to unintentional copyright infringement. Infringing users may be unaware that their use of a copyrighted work (e.g., stock photos in the case of Luo and Mortimer (2016)) constitutes infringement, they may be unaware of how to obtain a license for a work they found online, or the infringement may be a consequence of actions by a third party. For example, a firm may directly contract with a third party that then incorporates a copyrighted work into the firm's product without obtaining the necessary licenses. But in the digital era, there is also the risk that copyrighted works, or pieces thereof, may be added to databases without an appropriate license for such use. If these databases are then continually drawn upon to create new products and works, widespread unintentional copyright infringement could result.

Who are the targets?

Empirical evidence suggests that patent-owning NPEs preferentially target cash rich firms (Cohen et al., 2014). Rather than exclusively targeting practicing firms, the prevalence of digital piracy has led to copyright trolls also focusing upon anonymous consumers for statutory dam-

ages and settlements (Sag, 2015).[5] When firms and content creators are targeted by asserting firms, it is an open empirical question as to whether cash rich entities are the preferred targets. Anecdotal evidence, however, suggests that this may be the case. The independent label Illegal Art bills itself as "pushing the limits of sample-based music since 1998," disregarding the potential copyright infringements incurred in its distribution of sample-heavy albums that each may contain unlicensed use of several hundred songs.[6] Despite Illegal Art's disregard for the copyright implications of its releases, it has apparently faced no litigation for infringement. On the other hand, highly successful rap artists such as Jay Z, Kanye West, and Drake appear to often face litigation due to allegations of copyright infringement in their music.[7]

Do damages regimes matter?

Current damages regimes in copyright and patent policy incentivize litigation by non-practicing entities. Both patent and copyright policy obviate a need for plaintiffs to prove that infringement by the defendant led to realized lost profits in the plaintiff's market. This is an essential element for many IP intermediaries, as these entities may derive no revenue from their property rights absent that from litigation, settlements, and licensing royalties.

The owner or exclusive licensee of a copyright has the option of seeking actual damages from infringement or, as 17 U.S. Code § 504 stipulates, plaintiffs may instead seek statutory damages. When statutory damages are sought by a rightsholder, these awards range from $750 up to $150,000 for willful infringement. Statutory damages reduce the burden of proof on the plaintiff, and also may allow asserting firms to inflate settlements to an unreasonable size compared to the level of infringement (Sag, 2015). On the other hand, seeking actual damages may be quite lucrative with respect to infringement of music copyrights as discussed in this article, because the infringed rightsholder may seek the related profits from album sales, digital distribution, publishing revenue, concert ticket sales, and so forth.

---

[5]i.e., before the user record is subpoenaed from their ISP, the target and their ability to pay a large settlement is unknown.

[6]http://illegal-art.net/allday/samples.html (Accessed 5/3/2018)

[7]See, for example,
https://nyti.ms/2sQCZm1
https://www.spin.com/2018/02/kanye-west-solange-copyright-infringement-report/
http://www.ipwatchdog.com/2017/06/17/drakes-fair-use-copyright-victory-music-copyright-infringement/id=84504/
(All accessed 4/30/2018)

Just as copyright litigation may be incentivized by the prospect of statutory damages, PAEs need not prove to a court that lost profits were realized from the infringing use. As many PAEs are non-practicing entities, they may in fact derive no revenue from commercializing their patents aside from licensing revenue. Instead, PAEs may seek a reasonable royalty as damages for the infringement of their IP, assuming that their patents are valid and infringed. In contrast to the copyright regime in the U.S., plaintiffs in a patent suit may not request injunctions against infringement, or actual lost profits.

Scope of Rights

Patents in practice do not guarantee a well-defined property right to exclude infringing uses, but instead provide the right for an owner to attempt exclusion (Lemley and Shapiro, 2005). Additionally, patent litigation and assertion entities appear less prevalent in industries in which patent rights are typically not "fuzzy" and property rights have well defined boundaries – especially the chemical and pharmaceutical industries (Bessen and Meurer, 2008). In the copyright space, assertion and litigation may concern cases of explicit copying in whole - e.g., digital piracy of music and movies. However, the right to exclude is less well defined when a copyrighted work is incorporated into a new derivative work, or where such use may be considered fair use.[8] Copyrights may also be successfully asserted even in cases when a prior work is indirectly imitated or when a new work is merely similar to an old work.[9] Copyright infringement litigation in the U.S. thus often relies upon a standard of "substantial similarity" to determine whether a new work infringes upon the reproduction rights of a past copyright.

Despite this, the boundaries of copyright appear well-defined compared to patents. Copyright claimants cannot use vague claims as a strategy to increase the breadth of their rights, as can be achieved with patent claims. The invalidation of copyrights also appear to be rare.[10] Whereas a given technology may be covered by a patent bundle of uncertain size, the mapping between a creative work and its associated copyrights is very straightforward – any given work has a well defined bundle of rights defined by statute – though ownership and control of these

---

[8]See, for example, *Campbell v. Acuff-Rose Music, Inc.,* 510 U.S. 569 (1994) regarding 2 Live Crew's parody of Roy Orbison, or *Cariou v. Prince,* 714 F.3d 694 (2d Cir. 2013) regarding Richard Prince's appropriation of Patrick Cariou's photographs.

[9]See subconscious plagiarism in *Bright Tunes Music v. Harrisongs Music*, and the finding of substantial similarity between "Blurred Lines" and "Got to Give It Up" in *Williams v. Bridgeport Music*.

[10]Though we are not aware of statistical evidence so far that demonstrates the incidence of copyright invalidations.

copyrights can be a source of fragmentation.

Although copyright boundaries are well defined, the rightful ownership of a copyright, and hence standing to sue, can be murky given that copyright ownership is often fragmented among several parties. To bring legal action against copyright infringement in the U.S., the party must have an exclusive license to one of the aforementioned six exclusive rights that is allegedly infringed. Thus the asserting entity must have full ownership of the asserted right, not just a contract with the rights owner covering litigation of the owner or creator's copyrights. These issues have arisen in the litigation campaigns of copyright aggregators. In *Righthaven LLC v. Democratic Underground LLC*, Judge Roger L. Hunt determined that Righthaven lacked standing to sue for copyright infringement, as their agreement with the original rightsholder Stephens Media constituted a contract assigning standing to sue, not an exclusive license for exploitation of the litigated works. TufAmerica, an aggregator at the core of this article, faced similar problems asserting music copyrights against the Beastie Boys' sampling of songs originally recorded by Trouble Funk. The court determined that TufAmerica did not have standing in *TufAmerica, Inc. v. Diamond et al,* because while TufAmerica had exclusive licenses with two members of Trouble Funk covering all rights to their master recordings and musical compositions, TufAmerica had only an exclusive license to sue from the third remaining member of Trouble Funk. Without exclusive licenses covering exploitation from all three original copyright owners, TufAmerica did not own the copyright in question. Given this issue, it is an open question as to whether assertion entities preferentially acquire copyrights with unfragmented ownership, or rights for which exclusive licenses may be easily obtained.[11]

Intermediaries

The profitability of NPE's rests on their ability to act as arbitrageurs in the patent market. NPEs may purchase patents for low sums compared to the chance, albeit rare, that the NPE may use this patent to extract up to hundreds of millions in damage awards (Hagiu and Yoffie, 2013). On the copyright side, we are not aware of any work that demonstrates the typical or average cost to purchase copyrights by aggregators. Statutory damages for copyright infringement are of relatively small scale compared to the large sums possible with patent litigation, reaching a maximum of $150,000 per willful infringer. Seeking actual damages may

---

[11]e.g., works-for-hire owned by a corporation may be easier to purchase than a typical musical arrangement copyright with ownership spread among several artists or their heirs.

be a more lucrative prospect, given a plaintiff's ability to target all related revenue streams to the infringement, but it seems unlikely that such awards could compete in scale with the high end of potential patent damages.

The long duration of copyrights compared to patent rights likely forces the value of these property rights to depreciate at vastly different rates. Patent rights, with a term up to 20 years, rapidly depreciate in value. Copyright protection, on the other hand, may extend for 70 years after death of the author, or up to 120 years for works made for hire. Despite their long duration, the royalty streams from exploiting a copyright are likely to rapidly diminish due to consumers often preferring content of a recent vintage (Waldfogel, 2012). This may present an opportunity for aggregators to collect copyrights with high litigation value compared to the low practical value to the original creator.

Research on copyright litigation may help us better understand NPE's role as market intermediaries. The ability to identify pieces of copyrighted content in products could enable research to track whether NPEs solely enforce property rights on existing products, versus the extent to which they help license unproven rights.

3   Music Copyright and Sample Trolls

Digital sampling is the practice of re-using segments of prior sound recordings and musical compositions to create new sound recordings. Digital sampling was sporadically incorporated in music throughout the 20th century by experimental musicians, but saw widespread use in popular music with the commercialization of digital sampling devices in the 1980s, particularly in rap and hip hop music.

The U.S. Copyright Act of 1976 provides the current basis of copyright protection for musical compositions and phonorecords in the United States. Prior to 1991, unlicensed sampling of past recordings was rampant in hip hop music due to speculation that sampling may not infringe on the copyrights of the re-used musical arrangement and composition (McLeod and DiCola, 2011). In 1991, however, Biz Markie and Warner Bros. Records were sued by Grand Upright Music for the unlicensed sampling of Gilbert O'Sullivan's "Alone Again (Naturally)" on Markie's album *I Need a Haircut*.[12] The court's ruling against Markie and Warner Bros. was

---

[12] *Grand Upright Music, Ltd. v. Warner Bros. Records Inc.*

the first decision to recognize digital sampling as copyright infringement. In response, record labels were forced to scrutinize their artists' albums for unlicensed sampling, or else face the risks of costly copyright infringement litigation and the possible loss of all profits associated with the infringing use. The risk of litigation and the costs of clearing sample licenses led to a rapid shift away from heavy sample usage in the production style of hip hop music.

Since *Grand Upright*, split decisions have added uncertainty to the contours of copyright law with respect to sampling. In 2001, Bridgeport Music filed action against infringement of George Clinton's "Get Off Your Ass and Jam" in N.W.A's "100 Miles and Runnin"' regarding a two-second guitar riff that was sampled from the Clinton recording and re-used in the background of the song by N.W.A that was released in 1990. The subsequent 2005 decision of the 6th circuit in *Bridgeport Music, Inc. v. Dimension Films* determined that any duplication of a past sound recording, no matter the length, constituted copyright infringement, thus eliminating a defense of *de minimis* use with digital samples. However, in 2016, the 9th circuit rejected this "bright line" rule by the 6th circuit in *VMG Salsoul, LLC v. Madonna Louise Ciccone, et al.*, determining that any copying by Madonna of the plaintiff's song "Love Break" was *de minimis* and did not constitute copyright infringement.

The sampling and re-use of a past sound recording currently risks infringing two discrete copyrights of a previously published song, the copyright of the original sound recording as well as the copyright of the underlying musical composition. Music publishers, which typically control the composition rights of their songwriters, have historically been much more litigious over unlicensed sampling than record labels, which typically control the sound recording copy-rights of their artists. This difference in copyright enforcement may be due to the scope of publishers vs. labels. Publishers oftentimes specialize in a specific discipline or genre of music (e.g., jazz), compared to major record labels and their more general portfolio of artists, some of which may also be producing sample-based works.

The "clearance" of a license to sample a sound recording thus entails purchasing a license from two upstream monopolists - the owner of the musical composition copyright, as well as the owner of the sound recording right. This presents the familiar "complementary monopolies" problem, as each upstream monopolist may not consider the other's pricing decisions when pricing a license for the downstream use. In practice, sample-using artists oftentimes license only the composition right of a desired musical section, re-performing the licensed composition to create their own sound recording in order to to avoid the high costs associated with licensing both copyrights.

Even if a license is sought to clear a sample, a potential licensee may incur significant search costs in locating the appropriate rights holders. Copyright owners are under no obligation to register or record ownership changes in centralized databases like the U.S. Copyright Office's Public Records Database. This feature of the copyright system leads to the "orphan works" phenomenon for works whose ownership cannot be determined or whose owners cannot be contacted – an inefficiency that is greatly enhanced by the long term of copyrights, extending decades past a creator's death. The lack of a maintained database of registrations and transfers could also allow rights assertion entities to "ambush" infringing users who desired a license, could not identify rightsholders, and naively re-used a work.

Sample Trolls: Bridgeport Music and Tuff City

Bridgeport Music, Inc. was founded in 1969 by Armen Boladian, who also created the related Detroit based independent record labels Eastbound Records and Westbound Records. Westbound achieved moderate success in the 1970s with the signing of George Clinton's Funkadelic band as well as a brief record deal with The Ohio Players in between their stints at Capitol Records and Mercury Records. Samples of George Clinton, the Ohio Players, and similar funk musicians were instrumental in developing the sound of west coast hip hop music. In particular, samples of Clinton's Parliament and Funkadelic groups (affectionately known together as "p-funk") gained widespread sampling use in the late 1980s and early 1990s with the rise of the "g-funk" (gangsta-funk) sound.

On May 4, 2001, Bridgeport Music, Inc. filed approximately 500 claims of copyright infringement against all of the major music labels for rap and hip-hop songs that were alleged to contain unlicensed samples of copyrights controlled by Bridgeport Music. Many of the alleged infringing works, such as N.W.A.'s "100 Miles and Runnin," had been originally released during the early stages of hip-hop's history, over a decade prior to Bridgeport's litigation. This wide litigation campaign resulted in Bridgeport's popular categorization as a "sample troll."

Bridgeport and Boladian control the rights to the majority of George Clinton's work and his various related acts, including songs by Parliament, Funkadelic, Parlet, and the Brides of Funkenstein. The Bridgeport controlled works by Clinton were heavily sampled well before Boladian's litigation campaign, and many of Clinton's songs have been individually re-used and sampled hundreds of times, including Clinton's "Atomic Dog," a song that has seen re-use through almost 300 new songs. Since first observing the interest in re-purposing his works,

George Clinton appeared in public to be generally open and encouraging of sampling, such as through his release of "Sample Some Of Disc – Sample Some Of D.A.T," a CD that features sample-ready instrumental portions from a selection of previously-unreleased funk songs.

The control of the the rights to Clinton's early work has been subject to great dispute. Clinton denies the veracity of a March 1982 contract and a subsequent December 1983 addendum to the contract that transferred copyrights interests from Malbiz Music, Clinton's music publishing arm, to Boladian and Bridgeport Music. Clinton has claimed in court that these agreements were materially altered by Boladian and that his signature was forged. Since then, these agreements have also been featured in Clinton's bankruptcy proceedings as well as disputes over a Michigan farm that Clinton purchased in 1980. Although Clinton has achieved some success in reclaiming master recording rights to a few of his albums, the majority of the P-funk publishing catalog is still controlled by Bridgeport Music.[13] Bridgeport's continued ownership of Clinton's catalog is almost certainly short lived, as Clinton has since filed terminations under 17 U.S. Code § 203, actions that will revert the contested copyrights to Clinton over the next several years.

Tuff City Records is an independent New York City record label founded by Aaron Fuchs in 1981. Fuchs originally focused on early New York hip-hop music, releasing albums in the early 1980s by Spoonie Gee and The Cold Crush Brothers. Since then, Fuchs has shifted focus, and today bills the label as "rescuing thousands of blues, jazz, funk, soul and R&B treasures from obscurity."[14] Tuff City acquired this catalog of blues, jazz, funk, soul, and R&B over time, most notably with their acquisition of The Honey Drippers' 1973 "Impeach The President," a song that has now been sampled in over 700 new sound recordings. Fuchs has purchased licenses to build his copyright portfolio in other frequently sampled works, including songs by the band Trouble Funk, and attempted to litigate copyrights against popular artists like the Beastie Boys and Kanye West, albeit with mixed success.

While some have characterized Tuff's acquisition and litigation of copyrights in past vintage music as opportunistic rent-seeking, Fuchs has denounced this claim, "I have to contend with this clown Davey D's assertion that I bought records with break beats so that I could sue people. Everything I did was based on the notion that hip-hop would last."[15]

---

[13]i.e., Bridgeport still controls the composition copyrights.

[14]http://www.tuffcity.com

[15]http://fatlacemagazine.com/2007/12/tuff-city-records-part-1-interview-with-aaron-fuchs/

4 Empirical Strategy

To estimate the effect of copyright litigation on the magnitude of re-use, we employ the following model:

$$Samples_{it} = \delta \; PostTreat_{it} + \sum \alpha_t year_t + \sum \beta_i song_i + \varepsilon_{it}$$

Where $Samples_{it}$ is a count of the number of newly released songs that sample song i in year t, $PostTreat_{it} = 1$ for songs that are under the influence of sample trolls (discussed below), α_t captures year fixed effects, β_i captures individual song fixed effects, and ε_{it} represents an idiosyncratic shock.

In an ideal experiment to study the effect of copyright enforcement, the econometrician would randomly segregate sampling-prone songs into a treatment group and control group. The econometrician would then transfer the randomly-selected treatment group of copyrights to NPEs, agents that actively enforce, litigate, and license copyrights under their control. The exploitation of copyrights in the treatment group could then be directly compared to the control group.

Contrary to the ideal experiment, we observe copyright trolls endogenously collecting a portfolio of songs over time. This presents two challenges to identification. First, the copyright trolls are likely selecting and acquiring songs due to the propensity with which they are sampled, as well as the extent to which these songs were used in past works. That is, copyright trolls are acquiring the most valuable copyrights with regards to potential profits from litigation and licensing. Second, the acquisition of these songs coincides with the timing of court decisions that shifted sampling practices in the industry – though copyright trolls are certainly an important piece of the mechanism that moved the sampling industry from unlicensed use toward the current licensing model.

Our empirical strategy also explores two treatment timing assumptions for the variable $PostTreat_{it}$. First, we set 2001 as the treatment year for songs in Tuff City and Bridgeport Music's portfolio. We use this treatment date as this was the year that Bridgeport Music began its litigation campaign by filing mass lawsuits against infringing uses in the music industry. In an alternative set of estimates, we set the treatment date according to the following rule:

$$TreatDate_{it} = \max\left(t_{\text{troll } j\text{'s first lawsuit}}, t_{\text{song } i \text{ acquired by troll } j}\right)$$

This rule accounts for the fact that the rights of Tuff City and Bridgeport were acquired over time, as well as the fact that Tuff City began asserting its portfolio before Bridgeport.

Matching Procedure

To deal with the trolls' endogenous selection of high-value copyrights, we employ a matched-sample approach, where each "treated" song owned by a copyright troll is matched to an observably similar "control" song.[16] Our matching procedure selects one control group song for each treated song (i.e., a one-to-one match) by randomly selecting a song from the same genre and vintage of the treated song and whose cumulative samples in year 1991 are as close as possible to the treated song's cumulative samples at the same point in time. A song's genre was coded using the WhoSampled.com genres: Hip-Hop/Rap/R&B, Electronic/Dance, Rock/Pop, Soul/Funk/Disco, etc. Vintage was coarsened into half decade increments (e.g., 1990, 1995, 2000).

Upon matching each treated song to an observably similar control group song we maintain the assumption that the sampling trajectory of the treatment and control songs would not differ absent the ownership by the copyright trolls. This allows us to causally estimate the effect of copyright trolls on exploitation patterns.

5   Data

Sampling data was obtained from WhoSampled.com in February 2016. WhoSampled provides community-sourced data on the original sources for sample-based music, cover songs, and remixes. The raw WhoSampled data records information at the original song – sampling song dyad, including the year of release for both songs, the songs' associated labels of release, and the artists and producers for each song. This raw data also includes qualitative information on each sample, such as the portion of the original song that was sampled (e.g., drum break,

---

[16]Akin to the approach of Azoulay et al. (2010)

melody, vocals, etc.) and where the sample was re-used in the new song. Our sampling data includes 61,438 original songs by 21,688 original artists and 131,332 sampling songs by 32,250 sampling artists.

Data on the portfolios of Bridgeport Music and Tuff City Records was collected from several different sources. Bridgeport provides an online catalog of albums at their website, `bridgeportmusicinc.com`. All songs by these artists in the WhoSampled database were located in the online US Copyright office's registration and recordation database at `copyright.gov`. The earliest date on which a song's copyright was transferred to Bridgeport Music or Westbound Records was used as the date of acquisition. For songs that could not be found in the copyright office's database, the online BMI repertoire, `repertoire.bmi.com`, was searched to confirm that Bridgeport is the publisher for the copyright, and the song's original date of publication was used as the date of acquisition. Bridgeport's original song portfolio in our data is concentrated among just a few artists, with 166 songs in total owned by Bridgeport, 43 of these performed by Parliament, 47 by Funkadelic, and 15 by The Ohio Players.

Data on Tuff City's portfolio was collected by querying the copyright office's database for releases and recordations where Tuff City and/or TuffAmerica were named as a party. For all transfers between Tuff-labels and an artist, the first date at which a song appeared in a copyright transfer or publication agreement with Tuff was used as the date of Tuff's acquisition. Compared to Bridgeport, Tuff's portfolio is much less concentrated. Tuff has ownership interests in 49 songs in our data, with the most frequently observed artist, Trouble Funk, representing just seven of these songs.

Summary statistics for the estimation sample are shown in Table 1. These statistics come from the panel before one-to-one matching, therefore containing many "control" songs for each song in a copyright troll's portfolio. The sample is restricted to a window from 1987 to 2010.

5.1 Results

Selection Stage

Logit regressions exploring the selection of the copyright trolls' portfolios are shown in Tables 2 and 3, at years 1991 and 2001 respectively. It is apparent that both Bridgeport and Tuff are acquiring and retaining songs in their portfolio that have previously been subject to extensive use through digital sampling. Bridgeport Music's portfolio is heavily biased towards Soul /

15

Funk / Disco music due to their early acquisition of the George Clinton rights, apparent from the negative coefficients across genres when Soul is the omitted category.

The raw matched-panel trends in sampling of Bridgeport's songs are shown in Figure 1. Usage of Bridgeport's copyrights grows from the late 80s into the early 90s, coinciding with the g-funk style of hip-hop music popularized by Dr. Dre and associated acts on Death Row Records. By the late 1990s, the use of Bridgeport samples appears roughly comparable in magnitude to control group songs, but appears to drop further in the raw data after Bridgeport's litigation in 2001.

Analogous trends in the matched-panel data for the sampling of Tuff City's songs are shown in Figure 2. The trajectory for sampling of Tuff's music follows a comparable trend to Bridgeport's copyrights in Figure 1. Sampling of Tuff grows up until the early 1990s, before falling through the mid-1990s into the 2000s. In contrast to Bridgeport, Tuff's portfolio is less outlying than Bridgeport's, with the mean sampling of Tuff's portfolio actually below that of controls, at least until 1993.

5.1.1 Difference in differences analysis

Table 4 displays the results of differences-in-differences regressions on the full statistical sample in which the dependent variable, $samples_{it}$, counts the number of times a song was released in year t sampling song i. Columns (1) and (2) define the treatment year of copyright trolls' songs as 2001, the year that Bridgeport filed mass lawsuits against infringing labels. Columns (3) and (4) instead define the treatment date according to the following rule.

$$TreatDate_{it} = \max \left(t_{\text{troll } j\text{'s first lawsuit}}, t_{\text{song } i \text{ acquired by troll } j} \right)$$

All columns include year-level fixed effects. Column (1) includes the treatment variable of interest, $PostTreat_{it}$, which $= 1$ for songs in Bridgeport and Tuff's portfolio in years after 2001, as well as $Treated_i$, which $= 1$ for all songs owned by Bridgeport or Tuff. According to Column (1), trolls' songs on average are used in 0.64 fewer songs as a result of the 2001 mass litigation, but these songs on average have a higher baseline of sampling, with 0.74 more samples per song on average compared to control-group songs after controlling for year effects. The treatment effect from copyright trolls remains stable when song fixed effects are added in

16

Column (2), implying an average 0.63 decrease in the number of times song i is sampled, while also controlling for year-level fixed effects. Columns (3) and (4) vary the treatment date using the aforementioned rule, with Tuff's first copyright assertion lawsuit in 1991, Bridgeport's in 2001. In Column (3), the significance of the post-treatment measure drops out, while the positive and significant effect on $Treated_i$ remains. Column (4)'s estimate for the effect of copyright trolling is similar in magnitude to Column (2), but the estimate shrinks to 0.52 fewer samples per song.

A graphical depiction of these difference-in-differences results are shown in Figure 3. We observe an increase in the incidence of sampling for trolls' songs compared to control group songs leading up to the treatment date – here defined as the date of the *Grand Upright* ruling in 1991 – before a subsequent decline in the re-use rate for those songs in the trolls' portfolio compared to controls. However, given the clear pre-trend in the sampling trajectory for the treatment group compared to the control group, we should be concerned that the parallel-trend assumption of our difference-in-differences estimator is violated.

We further employ a matched panel approach due to the apparent imbalance between treated songs, those acquired by copyright trolls, and control group songs. Each treated song owned by Tuff City or Bridgeport Music is matched to a control group song by matching over song genre, vintage, and the nearest neighbor in terms of cumulative past samples at year 1991.

Logit regressions were run on the matched-panel in order to estimate selection-into-treatment. These results are shown in Table 5. In comparison to Table 3, we see that there is no significant difference in cumulative past samples for songs in the treatment group compared to the control group. When the sample is restricted to Tuff City (Column 2) or Bridgeport (Column 3), we still do not observe any significant imbalance on sampling between the treatment and control group songs.

Linear estimation results on a matched panel, where each treatment group song is matched to one control song, are shown in Table 6. All columns include year fixed effects. Columns (1) and (2) define year 2001 as the year of treatment, while Columns (3) and (4) vary the treatment date as previously described. In Column (1), we observe no statistically significant difference between treatment group and control group on the $Treatment_i$ variable, indicative of appropriate balance between treatment and controls. However, while the point estimate for the treatment effect translates to 0.12 fewer samples as a result of trolling, this effect is

insignificant at typical levels. This statistically insignificant treatment effect remains when song-level fixed effects are added in Column (2). No significant effects of copyright trolling are observed in Columns (3) and (4), implying that if trolls do exert a negative effect on the exploitation of works, any such effect is below the threshold of noise in our statistical sample. We also estimate the same specifications on the matched panel using poisson regressions, and observed no consistent statistically significant effect of copyright trolling with the results in Table 7.

Pre-trends and a graphical exploration of the matched-panel treatment effects are shown in Figure 4. In contrast to Figure 3, we observe no strong, statistically significant pre-treatment trend in the sampling trajectory of copyright trolls' songs compared to the matched control group songs. However, we also observe no statistically significant post-treatment change in the exploitation rate of the treated songs compared to controls. In fact, after treatment, the trolls' songs appear to have a slight increase in sampling according to the point estimates, but these estimates are insignificant at the 5% level, and these point estimates have attenuated to zero by five years post-treatment.

6   Conclusion

In this paper, we examine copyright trolls' acquisition and litigation of popular music copyrights in order to explore their impact on the supply-side use and re-use of digital samples. While it appears at first glance that these trolls exert a negative impact on the exploitation of a copyright under their control, these results do not remain once identification is carefully considered. Under our matched-panel approach that accounts for trolls selecting the most valuable songs, we cannot reject the null of no effect on re-use due to copyright assertion by the trolls. That is, when songs controlled by Bridgeport and Tuff are matched to similar control-group songs, we do not see any differential impact for those songs in the treatment group. However, this is not definitive evidence that assertions by Bridgeport and Tuff did not negatively affect the exploitation of these songs in the marketplace. While it is difficult to glean much information from this type of statistically null result, we can conclude that if there is an effect of trolling, it is not large enough in magnitude to distinguish it from the statistical noise in our sample.

Despite the statistical null result from the main troll treatment effect estimation, this study

does provide avenues for further research. Future work may further explore how copyright litigation has affected demand for inputs in music sampling. At a broader level, research could investigate the economics of copyright intermediaries – the cost of purchasing copyrights, success rates of litigation by intermediaries, what types of firms and producers get targeted with litigation, and estimates of average settlements and damages awarded from litigation.

References

Azoulay, P., J. S. Graff Zivin, and J. Wang (2010). Superstar extinction. The Quarterly Journal of Economics 125(2), 549–589.

Balganesh, S. (2013). The uneasy case against copyright trolls. Southern California Law Review 86(4), 723–782.

Bessen, J., J. Ford, and M. J. Meurer (2011). The private and social costs of patent trolls. Regulation 34(4), 26–35.

Bessen, J. and M. J. Meurer (2008). Patent failure: How judges, bureaucrats, and lawyers put innovators at risk. Princeton University Press.

Bessen, J. and M. J. Meurer (2014). The direct costs from npe disputes. Cornell Law Review 99(2), 387.

Cohen, L., U. Gurun, and S. D. Kominers (2014). Patent trolls: Evidence from targeted firms. Working Paper 20322, National Bureau of Economic Research.

Cohen, L., U. G. Gurun, and S. D. Kominers (2016). The growing problem of patent trolling. Science 352(6285), 521–522.

DeBriyn, J. (2012). Shedding light on copyright trolls: An analysis of mass copyright litigation in the age of statutory damages. UCLA Entertainment Law Review 19(1), 79–112.

Hagiu, A. and D. B. Yoffie (2013). The new patent intermediaries: platforms, defensive aggregators, and super-aggregators. Journal of Economic Perspectives 27(1), 45–66.

Lanjouw, J. O. and M. Schankerman (2004). Protecting intellectual property rights: Are small firms handicapped? The Journal of Law and Economics 47(1), 45–74.

Lemley, M. A. and D. A. Melamed (2013). Missing the forest for the trolls. Columbia Law Review 113(8), 2117–2190.

Lemley, M. A. and C. Shapiro (2005). Probabilistic patents. Journal of Economic Perspectives 19(2), 75–98.

Luo, H. and J. H. Mortimer (2016). Copyright infringement in the market for digital images. American Economic Review 106(5), 140–45.

McDonough, James F, I. (2006). The myth of the patent troll: An alternative view of the function of patent dealers in an idea economy. Emory Law Journal 56(1), 189.

McLeod, K. and P. DiCola (2011). Creative license: The law and culture of digital sampling. Duke University Press.

Sag, M. (2015). Copyright trolling, an empirical study. Iowa Law Review 100(3), 1105–1148.

Scott-Morton, F. and C. Shapiro (2016). Patent assertions: Are we any closer to aligning reward to contribution? Innovation Policy and the Economy 16, 89–133.

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

Figures and Tables

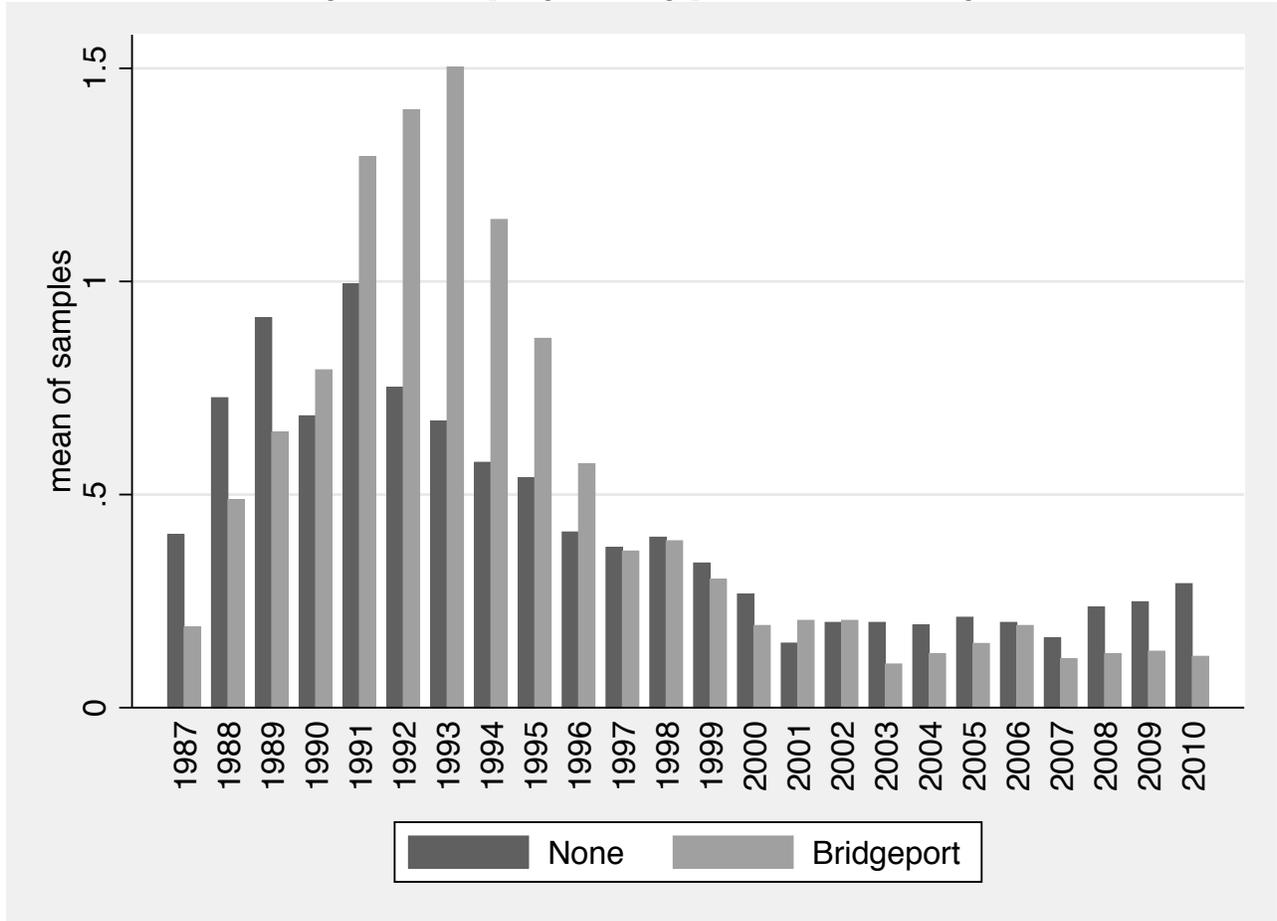Figure 1: Sampling of Bridgeport vs Control Songs

Figure 2: Sampling of Tuff City vs Control Songs
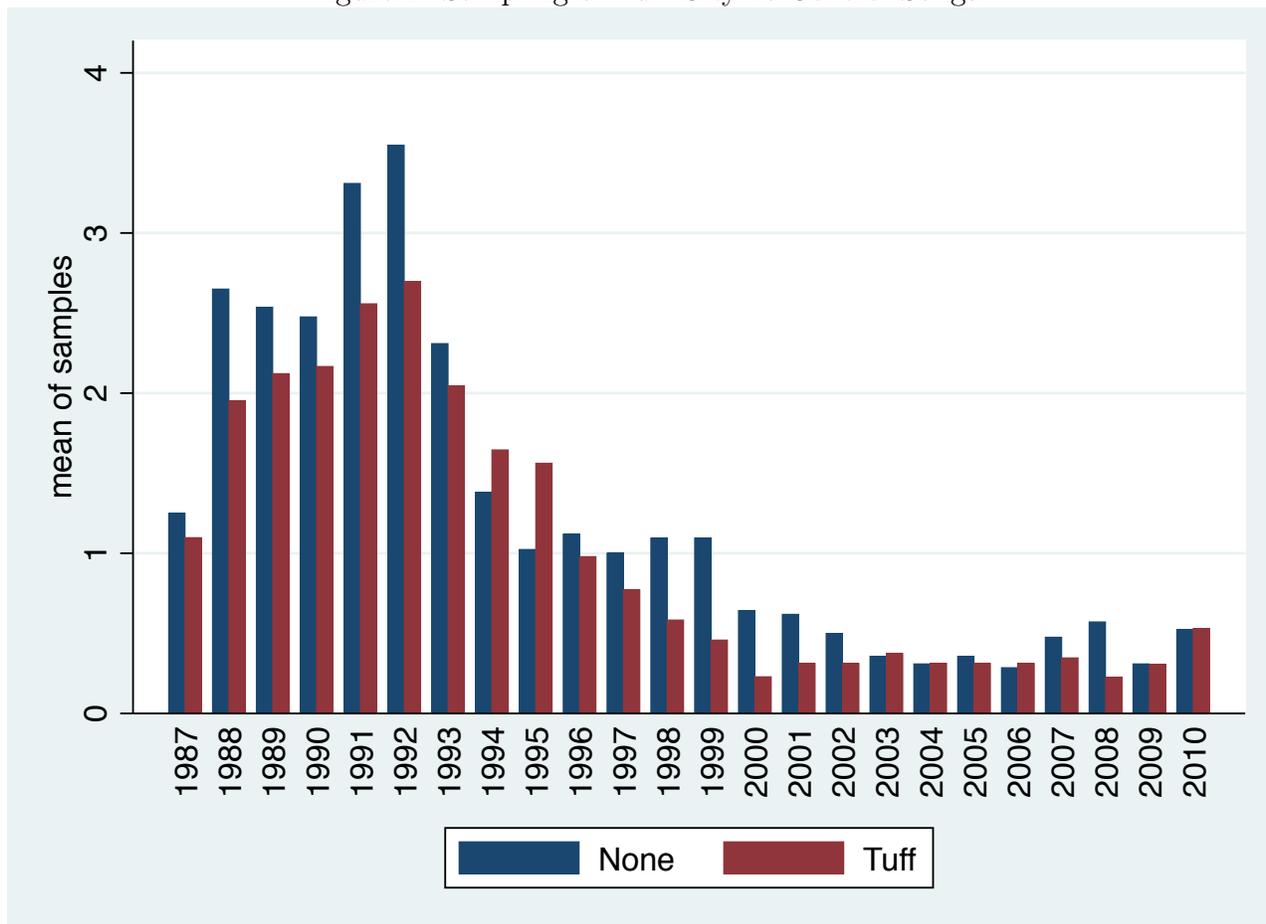
Figure 3: Pre and Post Copyright Troll Effects: Unmatched Sample
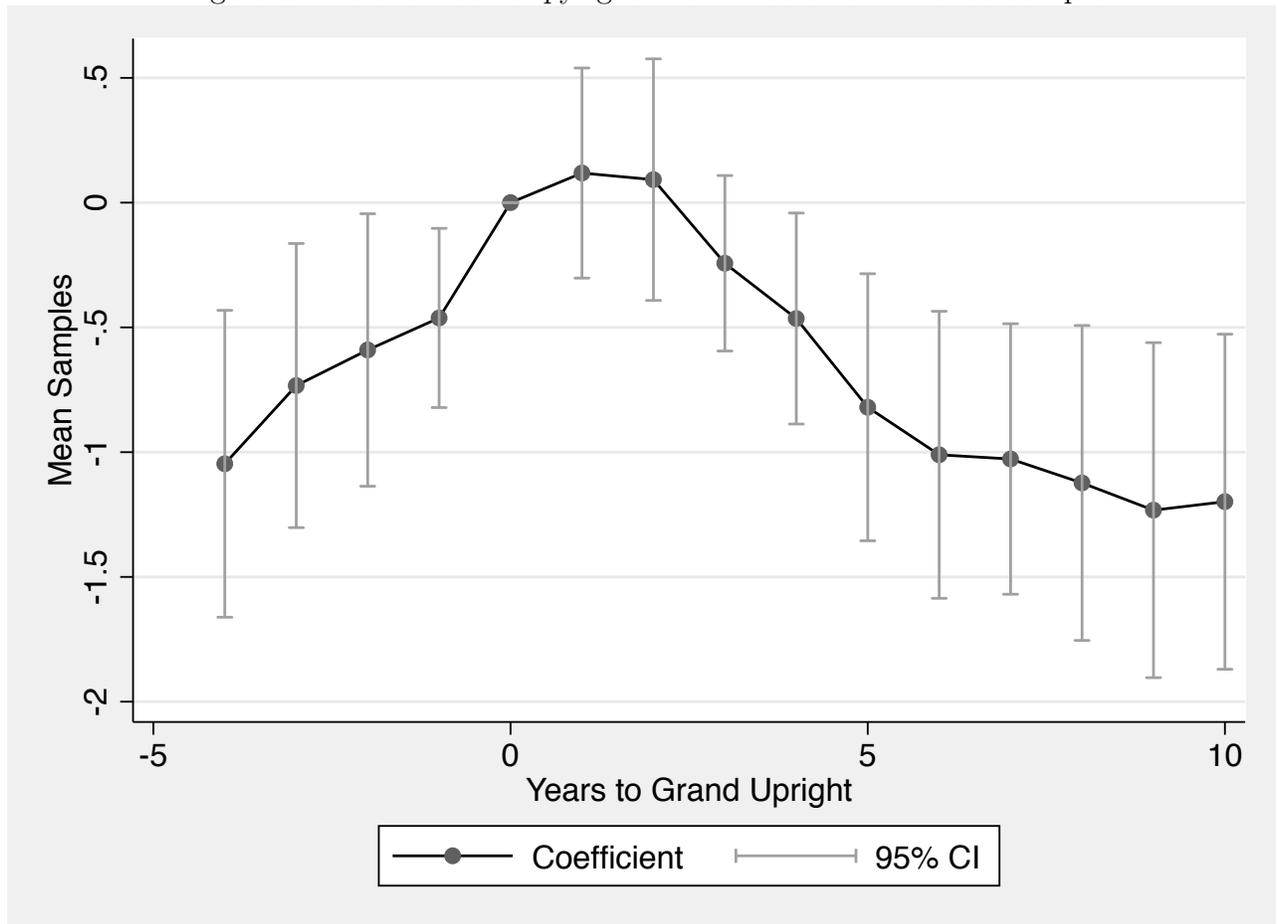
Figure 4: Pre and Post Copyright Troll Effects: Matched Sample
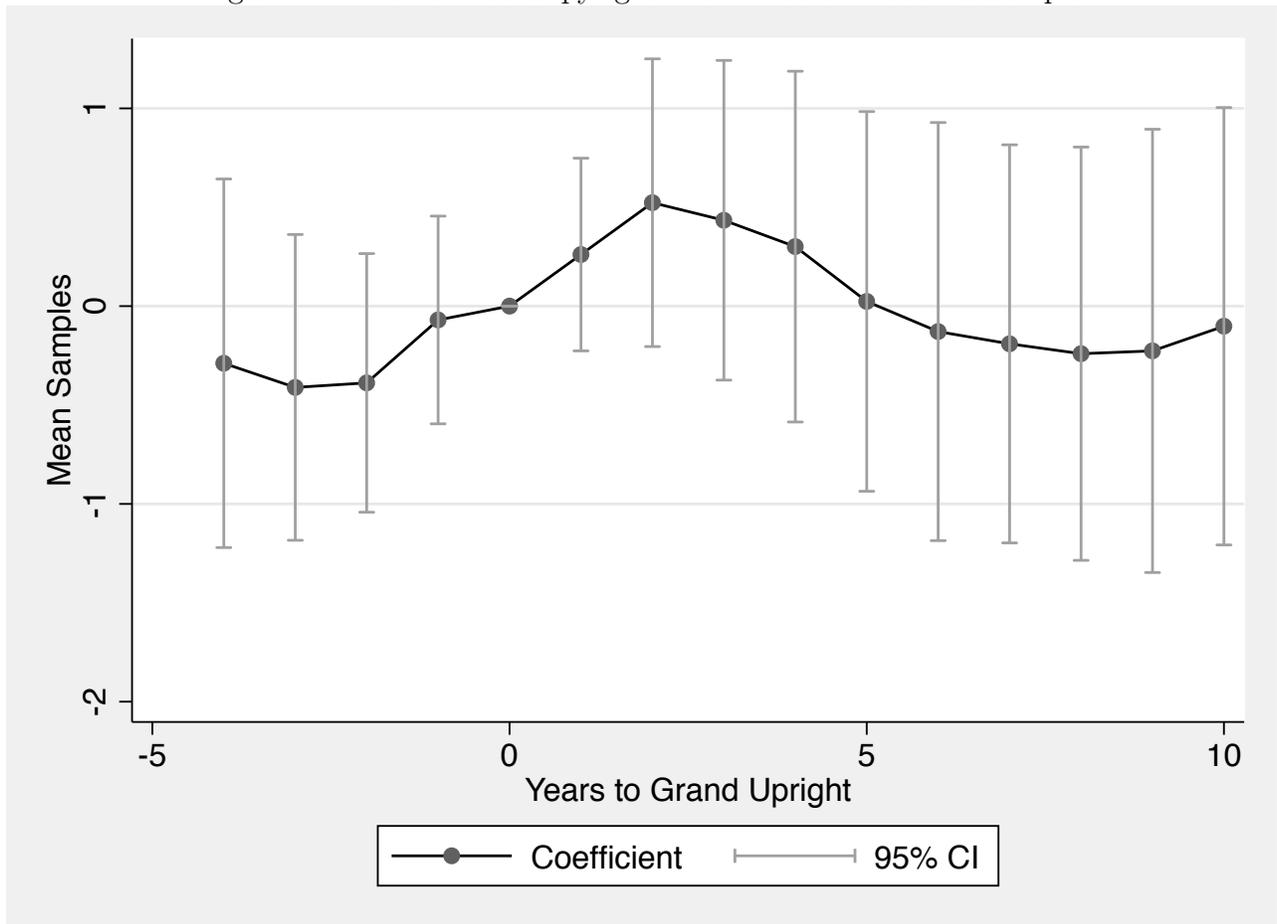
Table 1: Summary Statistics: Estimation Panel

	Mean	S.D.	Min	Max
Year	1999.64	6.80	1987	2010
Age	19.53	10.55	1	41
$Samples_{it}$	0.14	1.07	0	195
$Total\ Samples_i$	1.95	13.46	0	1615
$Troll_i$	0.005	0.07	0	1

Note: This table displays summary statistics from the pre-matching estimation panel

Table 2: Selection Regression: Year 1991

| | Logit | | |
| | (All) | (Tuff City) | (Bridgeport) |
	$Troll_{i,1991}$	$Troll_{i,1991}$	$Troll_{i,1991}$
$ln(Cumulative\ Samples_{i,1991})$	0.54**	0.58**	0.52**
	(0.06)	(0.11)	(0.07)
$ln(Age_{i,1991})$	0.31	0.68	0.14
	(0.21)	(0.39)	(0.24)
Hip-Hop / R&B	2.46*	2.42*	-3.80**
	(1.03)	(1.04)	(1.04)
Jazz / Blues	-0.68		-3.93**
	(1.43)		(1.00)
Other	0.02		-3.25**
	(1.13)		(0.51)
Rock / Pop	-0.58		-3.87**
	(1.24)		(0.71)
Soul / Funk / Disco	3.41**	1.15	
	(1.02)	(1.07)	
Observations	35,822	16,011	33,526

Standard errors in parentheses

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

Note: This table displays regression coefficients from regressions in which the dependent variable is is a binary variable which equals one if song i is controlled by a copyright troll in year 1991. Logistic regressions are used in each column.

Table 3: Selection Regression: Year 2001

	Logit		
	(All)	(Tuff City)	(Bridgeport)
	$Troll_{i,2001}$	$Troll_{i,2001}$	$Troll_{i,2001}$
$ln(Cumulative\ Samples_{i,2001})$	0.50**	0.45**	0.50**
	(0.05)	(0.11)	(0.06)
$ln(Age_{i,2001})$	0.75**	1.47**	0.34
	(0.27)	(0.48)	(0.31)
Hip-Hop / R&B	2.37*	2.43*	-3.79**
	(1.03)	(1.03)	(0.77)
Jazz / Blues	-0.56		-3.98**
	(1.43)		(1.00)
Other	0.43		-3.06**
	(1.11)		(0.46)
Rock / Pop	-0.37		-3.87**
	(1.23)		(0.71)
Soul / Funk / Disco	3.61**	1.31	
	(1.02)	(1.05)	
Observations	47,083	23,795	42,848

Standard errors in parentheses

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

Note: This table displays regression coefficients from regressions in which the dependent variable is is a binary variable which equals one if song i is controlled by a copyright troll in year 2001. Logistic regressions are used in each column.

Table 4: Baseline Regression: Unmatched sampled

	(1) $samples_{it}$	(2) $samples_{it}$	(3) $samples_{it}$	(4) $samples_{it}$
$PostTreat_{it}$	-0.64**	-0.63**	-0.32	-0.52**
	(0.17)	(0.17)	(0.22)	(0.12)
$Treated_i$	0.71**		0.59**	
	(0.21)		(0.15)	
Song FE	No	Yes	No	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	1,054,340	1,054,340	1,054,340	1,054,340

Robust standard errors, clustered at the song level, in parentheses

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

Note: This table displays regression coefficients from regressions in which the dependent variable is the number of times song i was sampled in year t. Linear regressions are used in each column. The treatment date is varied between Columns 1-2 and Columns 3-4. In Columns 1-2, the treatment year is 2001. In Columns 3-4, the treatment date is calculated as $TreatDate_{it} = \max\left(t_{\text{troll } j\text{'s first lawsuit}}, t_{\text{song } i \text{ acquired by troll } j}\right)$. For Tuff, $t_{\text{troll } j\text{'s first lawsuit}} = 1991$, while for Bridgeport it is 2001.

Table 5: Matched Panel Balance: Year 2001

| | Logit | | |
| | (All) | (Tuff City) | (Bridgeport) |
	$Troll_{i,2001}$	$Troll_{i,2001}$	$Troll_{i,2001}$
$ln(Cumulative\ Samples_{i,2001})$	0.05	0.05	0.06
	(0.08)	(0.13)	(0.08)
$ln(Age_{i,2001})$	-0.76	-1.80*	-0.34
	(0.54)	(0.76)	(0.70)
Hip-Hop / R&B	0.16	0.03	-2.14**
	(1.45)	(1.46)	(0.81)
Jazz / Blues	0.42		0.26
	(2.02)		(1.42)
Other	0.34		0.33
	(1.57)		(0.72)
Rock / Pop	0.43		0.25
	(1.75)		(1.01)
Soul / Funk / Disco	0.30	-1.23	
	(1.43)	(1.45)	
Observations	421	248	372

Standard errors in parentheses

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

Note: This table displays regression coefficients from regressions in which the dependent variable is is a binary variable which equals one if song i is controlled by a copyright troll in year 2001. Logistic regressions are used in each column. The sample is restricted to the one-to-one matched panel of treated songs to control songs.

Table 6: Matched Panel Regression

	(1) $samples_{it}$	(2) $samples_{it}$	(3) $samples_{it}$	(4) $samples_{it}$
$PostTreat_{it}$	-0.12	-0.12	0.43	0.09
	(0.26)	(0.26)	(0.48)	(0.25)
$Treated_i$	0.06		-0.18	
	(0.31)		(0.27)	
Song FE	No	Yes	No	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	10,047	10,047	10,047	10,047

Robust standard errors, clustered at the song level, in parentheses

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

Note: This table displays regression coefficients from regressions in which the dependent variable is the number of times song i was sampled in year t. Linear regressions are used in each column. The sample is restricted to the one-to-one matched panel of treated songs to control songs. The treatment date is varied between Columns 1-2 and Columns 3-4. In Columns 1-2, the treatment year is 2001. In Columns 3-4, the treatment date is calculated as $TreatDate_{it} = \max\left(t_{\text{troll } j\text{'s first lawsuit}}, t_{\text{song } i \text{ acquired by troll } j}\right)$. For Tuff, $t_{\text{troll } j\text{'s first lawsuit}} = 1991$, while for Bridgeport it is 2001.

Table 7: Poisson Matched Panel Regression

	Poisson			
	(1)	(2)	(3)	(4)
	$samples_{it}$	$samples_{it}$	$samples_{it}$	$samples_{it}$
$PostTreat_{it}$	-0.36*	-0.35	1.03	-0.08
	(0.19)	(0.18)	(1.04)	(0.24)
$Treated_i$	0.07		-0.19	
	(0.36)		(0.31)	
Song FE	No	Yes	No	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	10,047	9,305	10,047	9,305

Robust standard errors, clustered at the song level, in parentheses

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

Note: This table displays regression coefficients from regressions in which the dependent variable is the number of times song i was sampled in year t. Poisson regressions are used in each column. The sample is restricted to the one-to-one matched panel of treated songs to control songs. The treatment date is varied between Columns 1-2 and Columns 3-4. In Columns 1-2, the treatment year is 2001. In Columns 3-4, the treatment date is calculated as $TreatDate_{it} = \max\left(t_{\text{troll } j\text{'s first lawsuit}}, t_{\text{song } i \text{ acquired by troll } j}\right)$. For Tuff, $t_{\text{troll } j\text{'s first lawsuit}} = 1991$, while for Bridgeport it is 2001.