G R A P E
Group for Research in APplied Economics

Friends or foes? A meta-analysis of the relationship between "online piracy'" and the sales of cultural goods

Michał Krawczk, Joanna Tyrowicz, Wojciech Hardy

Foundation of Admirers and Mavens of Economics
Group for Research in Applied Economics

Friends or foes? A meta-analysis of the relationship between "online piracy" and the sales of cultural goods

Michał Krawczyk
University of Warsaw

Joanna Tyrowicz
FAME|GRAPE,
University of Warsaw and IZA

Wojciech Hardy
University of Warsaw

Abstract

Economic theory posits that consumers' response to technological innovations in distribution of cultural goods may have both positive and negative effects on the sales of these goods. Access to unauthorized cultural content -- often referred to as ``online piracy'' -- may reduce demand for the authorized distribution through substitution or raise it (through complementarity and similar effects). The empirical evidence speaks to both positive and negative correlation between unauthorized distribution and authorized sales. We review and discuss the accomplishments in the field so far and provide a quantitative analysis of the empirical literature, the meta-analysis. While numerous and interesting measures and methods have been developed, estimating reliable \emph{causal} effect remains a challenge. There are also substantial differences between film and music industry. On the whole the literature fails to reject the null hypothesis of no effects on sales.

Keywords:

digital piracy, cultural goods, meta analysis

JEL Classification

C92, D63

Corresponding author

Joanna Tyrowicz, j.tyrowicz@grape.org.pl.

1   Introduction

This study focuses on the link between activity referred to as "online piracy" and the sales of cultural goods. Analyzing this relationship is problematic and at the same time interesting for three main reasons. First, while "online piracy" is a popular phenomenon, its effects on sales have received considerably less attention from the scientific community than from commercial or civic actors. Second, "online pirates" typically try to hide their activity, making data relatively scarce. Moreover, the technological progress in production and distribution cultural goods introduces substantial changes to the nature of authorized and unauthorized acquisition of cultural content.[1] Third, funding in the field is often provided by the creative industries, who generally show interest in studies confirming the damaging character of the unauthorized distribution.

The field has no clear consensus about the sign and the magnitude of the displacement effect from unauthorized consumption to authorized sales. On the one hand, careful readers will find at least a few articles arguing in favor of a positive effect of unauthorized digital consumption. They rely on a range of arguments such as sampling and exploration, "dipping the toe", increased revenues for complementary goods and services, etc. On the other hand, some studies argue forcefully that "online piracy" actually substitutes legitimate sales. Indeed, zero or low price competition from unauthorized distributors can effectively reduce demand for usually higher-priced CDs, MP3 files, DVDs, movie theater tickets, etc.

In addition to conflicting results, there is also a disparity in measurement. Indeed, even the measures of legitimate sales or proceeds from authorized sales are often unavailable or proxied (e.g. by store survival). Naturally, "online piracy" is rarely directly observable, which necessitates the use of indirect measures or proxies. As a matter of fact, the very term of "online piracy" – though commonly used – is highly imprecise. Intuitively, most people identify it with acquisition of content from unauthorized sources, which is why the focus is placed on consumer choice. However, there is wide discrepancy in how this type of action is perceived: from a rational response of a consumer to technological progress with contested legal status to a crime punishable by imprisonment. In fact, acquiring copyrighted content is legal in many countries regardless of whether the distribution was authorized or not. For example, in many European countries – e.g. Germany, Spain, and Poland – only unauthorized distribution violates the law,[2] while acquisition is either entirely legal, or merely a tort, subject to typically troublesome civil lawsuit. By contrast, in the United States, acquiring copyrighted content from unauthorized sources is considered a crime. Naturally, the legal status of "online piracy" may also change over time.

---

[1] A few decades ago unauthorized distribution involved copying vinyls or 16mm tapes, which required substantial fixed setup costs and subsequently high distribution costs. In the 1980s authorized distribution moved to video tapes and cassettes, reducing the costs of unauthorized "industry" distribution, but also making it much easier to copy for own use. The onset of the Internet and digitalized cultural content made file-sharing and unauthorized distribution synonyms. Finally, the broadband access makes distribution gradually less about physical items or even actual files and more about streaming. A chronological overview of these changes along with data on US market was given by Oberholzer-Gee and Strumpf (2010). The changes in the means of cultural goods consumption make it challenging to measure meaningfully authorized sales as well as unauthorized acquisition, let alone a reliable counter-factual for authorized sales.

[2] A notable exception is the so-called fair use, i.e. sharing cultural content with family or friends in a way analogous to lending or passing on non-digitalized cultural goods.

Defining "piracy" is difficult also due to technological progress. Within the last decade, more and more unauthorized distributors make content available through streaming, which implies that the "downloader" actually never acquires the file but rather experiences the content in real time, from a server often located outside the jurisdiction of countries where most copyright holders reside. These services differ in business models: some rely on commercials, others request the so-called transfer payment, finally some distributors expect a regular fee for maintaining the account; to avoid criminal charges most unauthorized distributors offer the actual content for free. These are models similar to the authorized distribution; as a result the final users often find it hard to distinguish between these two types of sources (Liebowitz 2013).

In addition to the measurement issues, there are also conceptual complexities. Typically one would interpret content distributed with and without authorization as substitutes if cross-elasticity of demand was positive. However, in the case of acquisition of content from unauthorized distributors, the (marginal) cost is typically zero. Consequently, even if data was available on per-good per-client basis, cross-elasticity could not be computed.[3] Moreover, there are endogeneity issues: a positive correlation between (a proxy of) "online piracy" and legitimate sales cannot be meaningfully interpreted, because typically more attractive content will be more frequently acquired from both authorized and unauthorized distributors. Similarly, a consumer with a greater taste for culture may demand more from both types of sources.

With these two paramount problems (popularity driving both legitimate sales and "piracy" as well as near-zero price for "pirated" content), the field has offered a variety of methodological solutions. While some of them remain controversial – see for example Oberholzer-Gee and Strumpf (2007) and the subsequent critique by Liebowitz (2017) – they are designed with the intention to discriminate between substitution effects and complementarity effects. A strand of research attempted to instrument for "piracy" with variables seemingly unrelated to legal sales or to identify the effects through event studies (such as the introduction of new laws or the shut down of major unauthorized distributors). To avoid some of the delineation issues, studies typically analyze fairly similar cultural goods (e.g. only music or movies). Some of the studies also try to narrow the analyzed topic geographically (e.g. analyzing sales within a university campus or in shops in certain districts). Methodologically, it seemed relevant to several authors to control for individual tastes for culture or to apply fixed effects estimates to obtain estimates conditional on the unobservable taste for culture. This proliferation of methods, subjects and identification strategies constitutes yet another motivation to engage in a meta-analysis.

Our objective in this paper is to provide a systematic and quantitative evaluation of the achievements in this field of economics. We analyze the available research papers with two major goals. First, we aim to shed some light on the developing standards in this strand of the literature, emphasizing the relevance of endogeneity bias and the need for a strong identification strategy. We review measurements and methods used in the literature with the objective to evaluate external

---

[3]One way to address this problem would be to replace the actual price of content acquired from unauthorized sources with a fine multiplied by the individually perceived probability of being forced to pay this fine. However, fines are either absent or rarely given in most countries, making it rational to have near-zero expected probability of a fine, in turn indicating little variation in the "price" of unauthorized content and rendering estimation of cross-elasticity nearly impossible.

validity of some positive and some negative results. Second, we test the contention that with more robust methods and more comprehensive measures, the actual substitution effect is strong and large. Popularity of "piracy" is, after all, yet another example of a concept referred to in the literature as an enforcement problem. The ability to protect one's property from unauthorized use is not a new topic in economics – its traces may be found in numerous applications of empirical institutional economics, conflict economics and endogenous growth theories to name just a few.

Our study is not the first one to look at this growing body of literature (see Oberholzer-Gee and Strumpf 2010, Liebowitz 2013, Danaher, Smith and Telang 2014, Liebowitz 2016). However, there are some aspects that differentiate this study from the earlier contributions. First, we provide a systematic overview. This implies that the basis for our review is a universe of earlier studies, rather than an arbitrary selection. In fact, in total, we have identified as many as 45 studies (with 425 estimates), which is substantially more than covered by any of the previous reviews. Second, unlike previous studies, we employ a meta-analysis methodology, i.e. quantify the size effects in the literature using objective statistical methods. Danaher, Smith and Telang (2014) argue that a vast majority of published papers demonstrate a negative effect of "piracy" on sales. Meanwhile, of the 40 studies we identify, 21 interpret their findings as evidence for substitution rather than complementarity effects. However, there are also studies which find the opposite, studies which demonstrate no significant results and studies which argue that the direction of the link depends on the type of content or analyzed sample. In addition, in most papers, at least some of the specifications were insignificant or had a sign opposite to the claimed result. Unlike reviews, a meta-analysis provides tools to understand these discrepancies.

Following the literature in the field, we focus on consumer behavior. Thus, the term *acquisition* will be a shorthand for the acquisition of cultural content from unauthorized sources (e.g. downloading, streaming, copying, etc.). Given the mixed legal character of acquisition from unauthorized sources, as well as differences in the ethical perception of this activity, we also use the term *piracy* in the interest of brevity. Whenever we speak of *distribution* without consent of copyright holders, it will be explicitly indicated. In the next section we discuss the insights from the available literature in a broader context of institutional economics and property rights protection. In section 3 we discuss the literature sampling procedure and discuss of the insights from the available estimates. After this exploratory analysis, in section 4 we move to the meta-regressions, assessing quantitatively the role of robustness for the conclusions from the literature. The policy and research implications of our study are discussed in the conclusions.

2 Theoretical insights

Consider a set-up proposed by Belleflamme (2016). Assume that with no technology for unauthorized acquisition each consumer has unit demand, provided that the price is lower than his valuation v. Assume a unit mass of consumers in each period, whose valuations are distributed uniformly on $[0, 1]$. The good can be delivered at (low) marginal cost c and there is non-trivial fixed cost F of developing it. Once it is developed, per-period sales are given by $q = 1 - p$, so to

maximize her profits of $q(p - c)$, the producer charges a monopoly price of $p = (1 + c)/2$, with $(1 - c)/2$ consumers buying. Of course, unless the discounted flow of resulting profits over the product's expected lifetime exceeds F, it is not even developed.

When technology for unauthorized acquisition at, say, zero marginal prize, becomes available, authorized sales are fully displaced by the unauthorized distribution. Coordination problem emerges in order to finance the fixed cost F, but once this coordination problem is overcome, all consumers whose valuation v is above c consume the cultural good. Note that even in this simplified and admittedly extreme case, displacement is not as large as the unauthorized consumption, because – given downward sloping demand curve faced by an individual company, consumption at zero price will be larger than would have been under optimal pricing absent technology for unauthorized acquisition.

This simplest framework may be easily adjusted to account for imperfect substitution. Indeed, pirated content may have lower quality; it may come without physical complements such as printed materials; etc. Even more fundamentally, while e.g. digital files downloaded to a computer might be a good substitute for a DVD, this will not be true for a movie screened in a theater (especially if its appeal largely lies in special effects). If the pirated version is not a perfect substitute, positive authorized sales are possible at a positive price. Consider that the unauthorized copy may have lower perceived quality, $(1 - a)v$, with $0 < a < 1$. The consumer then only purchases (rather than acquires the pirated version) if $(1 - a)v \leq v - p$ or $v \geq p/a$. If the producer keeps the old price, she only sells $1 - (1 + c)/2a$. The remaining $(1 + c)/2a$ potential consumers acquire the unauthorized copy, so that displacement rate is $1 - a$. Displacement is large when the loss of quality in the unauthorized copies, a is small.

Note, however, that price $p = (1 + c)/2$ is no longer optimal. An optimal response on part of producers will typically involve lowering of prices. This is because availability of cheap substitutes generally leads to greater demand elasticity, so that profit-maximizing mark-up is lower. In our simple model, as the demand is now given by $q = 1 - p/a$, the producer should lower her price to $(a + c)/2$, yielding the sales of $(a - c)/2a$. The remaining $(a + c)/2a$ consumers will acquire the unauthorized copy, so that displacement rate will be $(c - ca)/(a + c)$. With lower price and quantity, the monopolists obtains lower profits. Displacement rate is low with high quality loss associated with unauthorized distribution; displacement will also tend to zero if c goes to zero. Depending on the demand function and other assumptions of the model, the reduced price due to piracy-induced price sensitivity could even result in increased authorized consumption, although not higher profits (see Bae and Choi 2006).[4]

Other authors account for possible positive impacts of piracy on sales. First, many digital products show substantial network externalities. If consumption value increases in the number of users, also unauthorized access may benefit legitimate users (Reavis-Conner and Rumelt 1991). For some distributions of taste for the original and pricing options available to the producer, this effect may even boost authorized sales. Second, potential benefit of unauthorized acquisition

---

[4]Other models consider a broader notion of utility loss, in particular associated with moral cost of unauthorized acquisition of the product (e.g. Chen and Png 2003). Likewise, Balestrino (2008) investigates the possible social norm of no-piracy; basing on functionalist view of norms.

involves sampling. Consumers who are uncertain about the value of the product can download an unauthorized version and, provided it turns out to match their needs, make an authorized purchase, see Gopal et al. (2006) for a single-product and Peitz and Waelbroeck (2006) for a multi-product setting. Both of these benefits of the low-cost version could in principle be provided by the producers themselves (King and Lampe 2003, Halmenschlager and Waelbroeck 2012, e.g. *freemium* models).

Finally, the literature argued that, at least in the case of music, online unauthorized distribution has increased the relative incentives for authors to create, because a file is not a substitute for a concert, but easier access to digital music makes consumers more inclined to participate in real-time events, because it lowers their costs of knowing new artists and new records (Waldfogel 2011, Liebowitz 2013). Naturally, the copyright industry does not benefit from growing popularity of live performances.

3   Literature sampling

We searched for empirical studies quantifying the relationship between online unauthorized acquisition and authorized sales of cultural content. We include both published and unpublished papers.[5] Likewise, we decided to include papers regardless of the notion of piracy involved in the estimation procedure. For example, some of the studies focus on exceptional events such as e.g. ABC temporarily disappearing from iTunes or shutting down of Napster and later MegaUpload. Narrowing analysis to these events only is likely to yield results of limited external validity and reliability. First, these events were outcomes of stakeholders' deliberate actions rather than random events, thus, they constitute poor natural experiments. Second, they were fairly rare.

We run the literature search in January of 2014 and we consider this the final date for the inclusion in our study.[6] Our starting point was the literature reviews by Dejean (2009), Grassmuck (2010), Smith and Telang (2012) which cite 27 relevant articles.[7] Additionally, a systematic search using EconLit data base was performed. Specifically, in EconLit we searched for (piracy AND sales) (piracy AND displacement) (piracy AND effect), ("digital piracy"), ("online piracy"), ("music piracy"), ("film piracy"), (sales OR revenue OR "box office").[8] To make our sample comprehensive, we also employed Google Scholar service with the same key words.[9]

In total our search procedure returned 73 research papers on the relationship between online

---

[5] While there are pros and cons of including "grey literature" (Higgins et al. 2008), most authors agree that the pros prevail (Tetzlaff et al. 2006). In our specifications we control for whether an estimate comes from a published paper or a working paper.

[6] If there was a working paper available in 2014, but the study was published later on, our database includes the published version.

[7] Given its later dissemination, we could not take follow the recent overview by Liebowitz (2016), but all of the studies he included in his overview are also included in our sample.

[8] The choice of the search words increased the selection of studies relative to previous reviews. For example Liebowitz (2013) includes in his review those studies, which estimate a *negative* effect of Napster shutdown on sales. Danaher, Smith and Telang (2014) focus on studies published in what they identify as tier journals.

[9] We are aware of the criticism raised against employing Google Scholar service for the search of material for meta-analysis. We made sure that some of the hazards associated with this method are minimized. First, we sought a published version, if it was available online but missing from EconLit. Second, we referred to the newest available working paper if the published paper was non-existent. Relevant articles were generally very scarce beyond the first 100 hits and we restricted our attention to the first 200 hits per search in the case of Google Scholar.

piracy and sales. We inspected all of the studies returned in response to this search. A share of articles were legal papers and philosophical essays, containing no empirical estimates of the relationship in question. Also, some papers were economic, but purely theoretical or descriptive. In addition, some studies discussed the drivers of the downloading behavior rather than the overall effects on sales () papers or articles that did not *demonstrate* the analysis of online piracy and sales could not be considered. This second group of articles and working papers includes some relatively highly cited ones, the authors of which e.g. assumed a degree of harm done and calculated the 'actual' losses. Overall, we kept papers that contain original empirical work quantifying the link between some measure of piracy and some measure of legal sales of the music or films.[10] We could only include papers that contained coefficients on explanatory variables identified as (in)direct proxies for piracy and explained variables identified as (in)direct proxies for sales. In the third step, after careful inspection, the analysis was narrowed to the papers that actually estimated the effects of online piracy on sales, rather than other relationships. For example, some authors study the proportion between authorized and unauthorized content acquisition (Sandulli 2007) or overall demand for music consumption (Aguiar and Waldfogel 2018).[11] This procedure reduced the final sample for analysis from 73 to 45 studies. To set this number in perspective, Danaher, Smith and Telang (2014) include 18 studies in their review, whereas the review by Oberholzer-Gee and Strumpf (2010) comprises 14 studies.

The final sample of 29 published articles and 16 working papers yields 81 general findings about the effects of unauthorized acquisition on sales (some studies analyzed more than one type of good, or considered different samples) and 425 estimates that could be studied through a meta-regression, see Table A.2; the list of all papers used in our study can be found in Table A.1 in Appendix A. The sample is fairly balanced between studies on movie and music industry. The former are relatively abundant in recent years. The findings for the film industry seem more consistent, although in both cases there is support to both positive and negative effects of online piracy on sales, see Table 1. One interpretation is that in the first years of the 21st century Internet piracy was considered a threat mainly to the music industry and was therefore studied more vigorously. Indeed, it is the music-sharing platform Napster that is often referred to as the starting point of the piracy age. One should also remember that files with films are much larger and therefore more difficult to transfer – leading to the conclusion that film piracy might have reached a higher scale only subsequent to the increasing coverage of broadband internet and reduced fares for access.

3.1 Descriptive statistics

We start with identifying the general finding of each study concerning the nature of the relationship between unauthorized acquisition and authorized sales: substitution, complementarity or inconclusive, see Table 1. This is based on the authors' statements in the abstract, introductory, or concluding section, the procedure involving as little discretion from our side as possible. The

---

[10]It should be noted that we have searched for studies in industries other than movie and music, although with poor outcomes - we have only identified one study on the impact of piracy on video games sales (Fukugawa 2011) and therefore we do not include it in any of the following tables and analyses.

[11]The list of exclusions is discussed at length in Appendix A.

negative effect is far from universal. In fact, as much as a third of conclusions concerning the film industry are of no effect, whereas approximately one in five conclusions on music are of a positive estimate of the link between piracy and sales.

Table 1: Relationship between unauthorized acquisition and sales according to the authors in the field

Year of publication	Film industry			Music industry		
	Substitution	Inconclusive	Complementarity	Substitution	Inconclusive	Complementarity
2004	-	-	-	2	2	3
2005	-	-	-	2	0	0
2006	2	1	0	5	0	2
2007	5	0	0	1	2	0
2008	-	-	-	2	0	1
2009	1	1	0	2	1	1
2010	1	1	0	3	1	0
2011	1	1	0	2	0	0
2012	2	3	1	2	0	0
2013	2	1	2	5	1	1
2014	1	1	0	9	3	3
Published	12	7	1	23	4	1
Working paper	3	2	2	10	6	10
Total	15	9	3	33	10	11

Note: classification per paper, based on the interpretation by the authors. Substitution implies a negative estimate of the coefficient measuring the correlation (or causal effect) of unauthorized acquisition or distribution on authorized sales. Complementarity implies a positive coefficient. Admittedly, the coefficient may be positive due to network externalities, sampling, etc. Hence the term complementarity should be considered as a shorthand for the universe of the potentially positive effects.

These conflicting results may stem from a number of effects. First, as mentioned before, in theory there could be both positive and negative effects of piracy on sales. If their sizes are comparable, the failure to reject the null hypothesis may actually mean *strong opposing effects* rather than *no effect*. Second, all the estimates may be biased due to the omitted variable problem, sample selection, as well as reverse causality. Not all studies attempt to account for this endogeneity, which necessitates caution in deriving conclusions from this table. Third, while majority of the studies used US data, the 45 analyzed research papers consisted also of studies for Australia, Canada, China, France, Germany, Japan and Sweden, along with several cross-country studies.

The final explanation for the differentiated results summarized in Table 1 consists of the methodological differences in how both online piracy and sales are measured or proxied. The studies of the film industry include analyses of piracy effect on box office revenues, DVD sales, TV watching, video rentals etc., while the ones of music industry distinguish between digital and physical formats of music. To simplify the categories we have constructed seven groups based on essential differences between the explained variables. Four of them are clearly related to consumer-level data on individual behaviur. These usually came from surveys or some forms of tracking software. The remaining three groups use producer or industry-level data, such as on general sales volumes, revenues or chart ranks, see the upper part of Table A.3.

The way that unauthorized digital acquisition (which is typically unobservable) was proxied for in the analyzed papers is rather heterogeneous as well. Indicators such as downloads or page visits may be informative for some sources only. Moreover, unauthorized distributors rarely make data on popularity rankings or downloads available, even for research purposes. Some studies thus

assume that rather than the actual act of online piracy, a proxy should focus on the 'window of opportunity'. Consequently, many authors rely on broadband/internet connection, relevant skills of the user, or access to means that can be used for online pirating. These studies should be clearly separated from those using institutional change as a proxy (such as a change in law or a shutdown of a major source of unauthorized content). Based on a careful examination, we have classified these predicting variables, see the bottom part of Table A.3.

Given the wide dispersion in measurement of authorized and unauthorized consumption, the dispersion of the conclusions in the literature becomes less a sign of controversy and more a sign of the heterogeneity of the analyzed phenomena. There is a number of reasons why the overall conclusion on the link between online piracy and sales could be perfectly consistent and yet differentiated across papers. First, music and film industry could react differently to online piracy, also because its scope increased in different time periods. Second, it is a fairly general phenomenon in economics that earlier studies are typically less concerned with the sample selection, omitted variables and endogeneity biases, whereas the later studies more often seek a clean causal identification strategy, which may happen at the expense of precision too. Finally, if both positive and negative effects are present, it is possible that one of them accidentally dominates in a particular context. Thus, accounting for industry specificity, method, and context of the study should reveal a pattern in the results. Table 2 summarizes.

Table 2: Estimates in the literature by methodology and industry.

	No. of papers	No. of estimates	% Positive		% Negative	
			All	Best	All	Best
Methods						
OLS/probit/tobit	23	230	38%	31%	45%	50%
Panel OLS/probit/tobit or ARMA	16	89	21%	22%	53%	54%
IV	11	96	30%	25%	27%	56%
PSM and similar	3	10	10%	0%	70%	50%
Industries						
Movie (DVD or VHS sales)	5	18	11%	15%	44%	38%
Movie (Digital)	1	12	0%	0%	100%	100%
Movie (TV watching)	2	42	40%	29%	29%	21%
Movie (Theatre)	8	64	19%	8%	56%	73%
Movie (DVD or VHS rental)	3	13	8%	0%	38%	33%
Movie (Other)	2	34	21%	30%	50%	55%
Music (Album sales)	15	144	22%	31%	44%	49%
Music (Song sales)	6	60	61%	48%	39%	52%
Music (Other)	4	38	76%	43%	21%	57%

The pattern emerging from Table 2 is partly blurred by the fact that some of the papers present multiple specifications, of which only some are considered valid by the authors (e.g. authors first show a biased OLS estimator to subsequently move to interpreting the unbiased estimator from an IV specification). In response to this tendency, for all the studies in our meta-regression sample we have identified regressions that could be considered 'reliable sets'. Some of the authors specifically pointed at regressions they considered to be appropriate for interpretation or that they favored. For these studies the choice of best sets was straightforward. For studies that reported only one result (or several but for different samples/types of good), we marked all the coefficients as 'reliable'.

For the remainder of the studies we employed the following procedure. We mark an estimate as 'reliable set' if:

- the author(s) followed a particular specification for the robustness checks (i.e. referred to a previous specification in the main study)

- the author(s) quoted, in the abstract or the conclusions section, the result(s) from particular specification(s),

- the author(s) quoted a particular result to compare effect sizes between robustness checks and previous specifications

- the author(s) performed a regression that they indicated will control for potential biases in previous regressions.

If the author(s) reported several regressions but made no indication on which they considered better, we have included them all as reliable. Given how substantial the endogeneity bias can be in the study of the effects of online piracy on sales, we place great emphasis on identifying studies which made provisions to address this issue. This procedure yielded 283 'reliable' estimates.

4 The meta-analysis

We first provide evidence on all analyzed papers (separately for music and film industries) using the forest plots. We consider the movie and music industries separately because technological progress in the online exchange of files is likely to affect these two industries differently. Moreover, there is a substantial difference between these two industries in how unauthorized acquisition is proxied and analyzed. The forest plot method is the most transparent in a sense that we impose no classification of the data. It is also particularly well suited given the dispersion of data and methods in this literature. In the forest plots the dashed vertical line shows the line of 0 size effects, i.e. in our case, no correlation between authorized sales and unauthorized acquisition. For each study, the average size effect is estimates, along with the standard error of this size effect. If the dispersion of the results is large, the box signifying the distribution of effect sizes in a given study encompasses 0, thus pointing to no robust correlation.

Figure 1 portrays that with two exceptions for the film industry and two exceptions for the music industry, all studies are in fact too under-powered to report unequivocal effect sizes. Naturally, it could be that the results are very robust and if the authors had the chance to report a wider variety of robustness checks, our estimation would produce smaller within-study standard errors and thus more precise effect sizes. However, at this point, the literature typically delivers mixed results, even if we narrow the analysis down to the those coefficients which authors themselves classify as reliable. In the case of film, in two cases within study variation of results is such that the overall conclusion from the paper is significant at all. While for the film, both are negative, for music industry one yields a positive effect size and one yields a negative effect size.

The funnel plots in Figure 2, indicate the scope for the so-called publication bias. The solid line in both plots indicates the mean of estimated effects, and the dashed lines that fan out from

the top of the funnel shows the 95% confidence area where most of the estimates would fall if the dispersion in estimates was driven solely by sampling error. Publication bias is indicated whenever a disproportionate number of estimates lie on one side of the inverted, V-shaped confidence area. Funnel plots suggest there is publication bias in favor of negative estimates for the film industry, but for the music industry there appear to be many estimates both "too" positive and "too" negative. Further, the wide dispersion at the top of the funnel is consistent with substantial heterogeneity in methods and measurement. These results confirm that the literature is inconclusive, with significant number of surprising estimates of both signs. While there is a larger number of papers (and estimates) emphasizing the strong sales displacement effects, this group of studies is in general neither large nor conclusive.

The comprehensive results from a meta-regression shed light on the earlier forest plots and funnel plots. Unlike forest plots, meta-regression has more statistical power, because it exploits the fact that many different studies are considered at the same time. Unlike funnel plots, meta-regression allows to account for part of the heterogeneity within and between studies, thanks to the inclusion of controls for the studies characteristics. We choose to apply the random effects as our sample clearly includes estimates that are not easily comparable: the sample varies in the applied methodologies, selected samples, and even the nature of the studied goods. We also report the results for the weighted least squares (WLS), following suggestions of Stanley and Doucouliagos (2015). Admittedly, advantage of WLS over the random-effects model is stronger in the presence of publication bias. The results for the movie and music industries are reported in Table 3. In text we discuss the estimates for the 'reliable' estimates, the full set of literature estimates is reported in Appendix B.

Once we account for the characteristics of the data and the methodology, we find that in fact the estimate of a constant across the studies is typically insignificant for both film and music industry, as reported by columns (1)-(6). This finding is consistent with Liebowitz (2013) argues that the overall support to the substitution hypothesis in the literature is weak. The constant is insignificant also for the studies published in a journal with an impact factor, but once we include impact factor of the outlet among the controls, the constant becomes large and negative. This finding is consistent with the reviews of Oberholzer-Gee and Strumpf (2010), Danaher, Smith and Telang (2014), but our collection of analyzed studies is far more comprehensive.

Overall, our results speak to strong publication bias towards the substitution effect, once controlling for the impact factor of the journal. However, this need not signify mechanisms undermining the credibility of economic research in the field: given the small number of causally identified studies, it could be that the best published studies simply happen to identify larger displacement effects whereas studies with weaker casual identification (which automatically find smaller effect sizes due to the endogeneity bias) tend to be published in lower ranked journals. This interpretation is consistent with the findings in columns (7)-(12), where we report analogous results for the inverse standard error: once we control for impact factor of the journal, published studies display a tendency that more precise estimates point towards substitution effect rather than complementarity. Interestingly, ,there are differences between film and music industry: the higher the

impact factor, the more likely the substitution effect in the case of films, and the more likely the complementarity effect in the case of music.

The studies in the literature report the effects from early 1990s until mid 2010s. Over this period, many technological changes occurred. In the early years, unauthorized acquisition was predominantly about acquiring substitutes of the CDs or VHS/DVDs. Growing speed and availability of data transmission as well as the introduction of many alternative consumption models for cultural goods – music and film – changed remarkably the consumption patterns away from traditional mediums, towards computers and mobile devices. Likely, the nature of unauthorized acquisition has changed, possibly altering the relationship between piracy and sales. We inspect this hypothesis by including controls for the period of study: beginning and the end of the sample period. We find no time effects for the film industry, but there appear to be a robust tendency towards substitution effects with the time for the music industry. These effects are economically small, but economically robust. Consequently, if anything, the unauthorized and legal copies were becoming become less substitutive and more complementary over time.

These conclusions are confirmed by plots 3 and 4, where we report the number of studies for a given period in our sample, which report positive/negative effect size (according to the estimates reported as 'reliable' by the authors). These are unweighted averages, in a sense that if a study covered period 2005-2007 and reported substitution effect, we assume that this effect was homogeneous for 2005, 2006 and 2007. We then average the estimates from each covered period. For the figures, we exclude those years, for which less than three studies are available. In the interest of clarity, we do not display inconclusive results.

All our specifications control for the characteristics of the sales measurement/proxy and unauthorized acquisition measurement/proxy. We also control for the cross-country comparisons and physical consumption in the form of box-office. Finally, we control for whether or not the authors acknowledge the risk of endogeneity and if they claim they address this issue in their specifications. We make no discretionary judgment on whether this way of addressing endogeneity is sufficient. One would be tempted to expect the results which account for endogeneity to show smaller displacement effects than the results which do not account for this bias.[12] However, the estimates for all these controls not very robust. One of the reasons behind this lack of robustness is that a large number of dummy variables exists only in some configurations in the literature. For example, cross-country studies usually use OLS with country fixed effects and instrument or proxy for unauthorized acquisition with internet penetration. There is virtually no exception from this rule, which means that identifying separately the role of internet penetration as a proxy and cross-country studies as a sample is not feasible.

---

[12]This intuition need not be warranted. For example, Blackburn (2006) argues that unauthorized distribution helps the sales of less popular goods and does harm only to the highly popular ones.

Figure 1: Forest plot for film studies

(a) film studies

(b) music studies
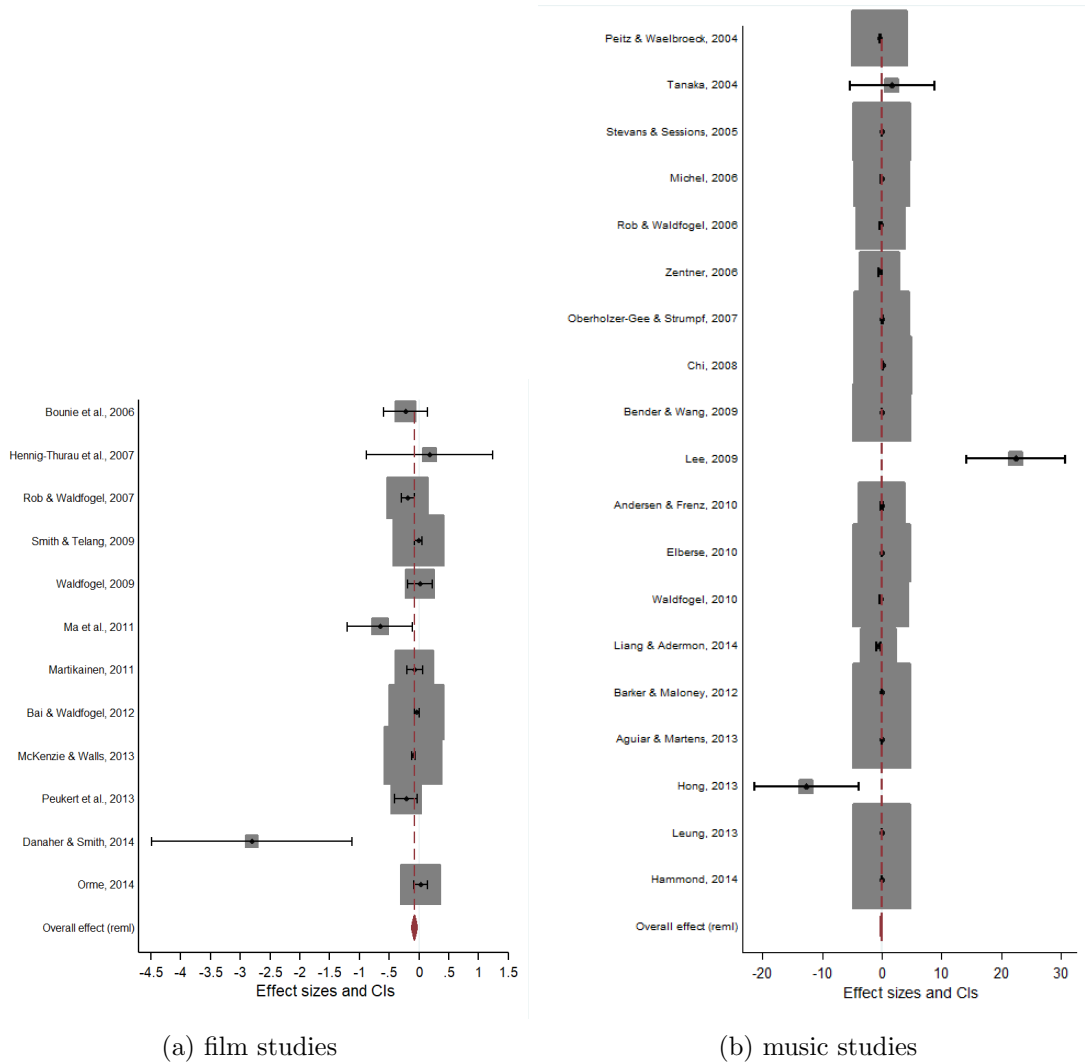
Note: Random effects maximum likelihood estimator, on a subsample of 283 estimates classified by the authors of the studies as 'reliable'. Paper by Blackburn (2006) first shows regressions with proxies of "online piracy" on RHS and subsequently proceeds to interactions of this variable without reporting own effect. The estimates on the interaction terms are considered 'preferred coefficients' by the author, but cannot be included in this forest plot for the lack of comparability.

Figure 2: Funnel plots

(a) film, 'reliable' coefficients

(b) music, 'reliable' coefficients
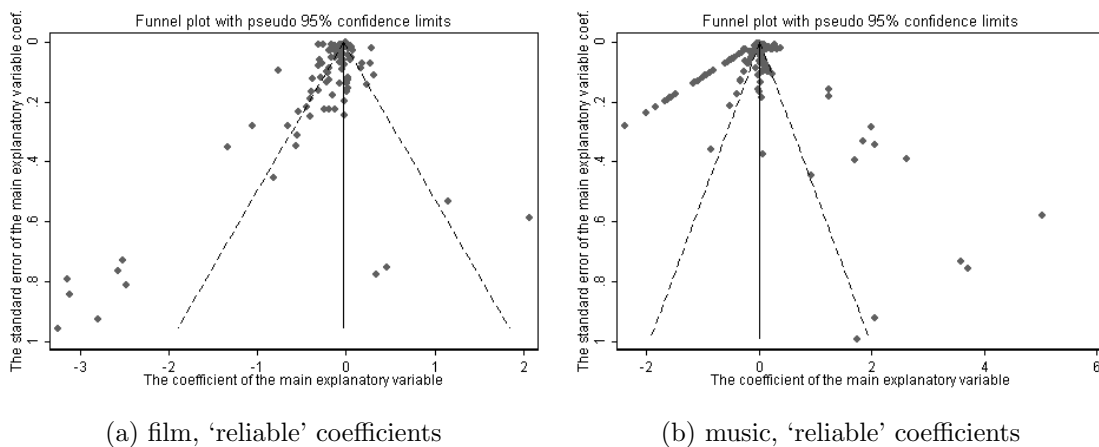
Note: Funnel plots for a subsample of 283 estimates classified by the authors as 'reliable'.

Table 3: Regressions for the film industry

	RE						WLS					
	Film			Music			Film			Music		
	ALL	ALL with IF		ALL	ALL with IF		ALL	ALL with IF		ALL	ALL with IF	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Constant	-10.502	-64.569	-157.895***	-16.552	-41.037	-158.267***						
	(26.241)	(50.141)	(56.059)	(16.067)	(30.668)	(26.975)						
Precision ($1/SE$)							20.144**	-5.714	-121.139***	-7.313	-18.112***	-124.958***
							(8.541)	(27.638)	(40.521)	(5.000)	(5.359)	(21.766)
Journal's impact factor			-0.564***			0.155***			-0.019			0.103***
			(0.191)			(0.026)			(0.086)			(0.020)
Year of publication	0.017	0.011	-0.107**	0.026**	0.114***	0.196***	0.003	-0.001	-0.005	0.000	0.077***	0.148***
	(0.013)	(0.014)	(0.042)	(0.012)	(0.040)	(0.031)	(0.009)	(0.006)	(0.019)	(0.004)	(0.019)	(0.022)
Period of study (first year)	-0.011	0.022	0.189***	-0.017**	-0.094***	-0.118***	-0.011	0.005	0.012	0.004*	-0.068***	-0.086***
	(0.011)	(0.025)	(0.061)	(0.008)	(0.030)	(0.021)	(0.008)	(0.011)	(0.035)	(0.002)	(0.017)	(0.015)
Study span (in years)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of variables used	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of the observations	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample selection acknowl.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Reverse causality acknowl.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE for sales proxies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cinema based study	Yes	Yes	Yes				Yes	Yes	Yes			
FE for "piracy" proxies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# of coefficients	103	73	73	179	76	76	102	72	72	159	75	75
R-squared							0.686	0.364	0.365	0.488	0.724	0.800

Notes: We report results for the coefficients which according to the authors of the studies were reliable. Table B.5 reports analogous estimates for all reported coefficients. Columns (1) to (6) obtained from `-metareg-` in STATA. Columns (7)-(12) show weighted least squares following Stanley and Doucouliagos (2015), with LHS and RHS variables divided by the standard error of the effect. Standard errors in parentheses.*** p<0.05, ** p<0.10, * p<0.15.

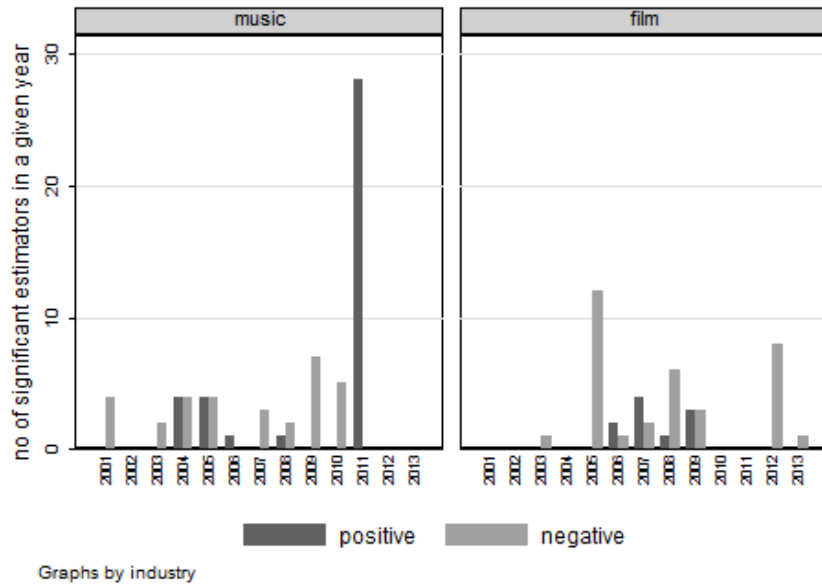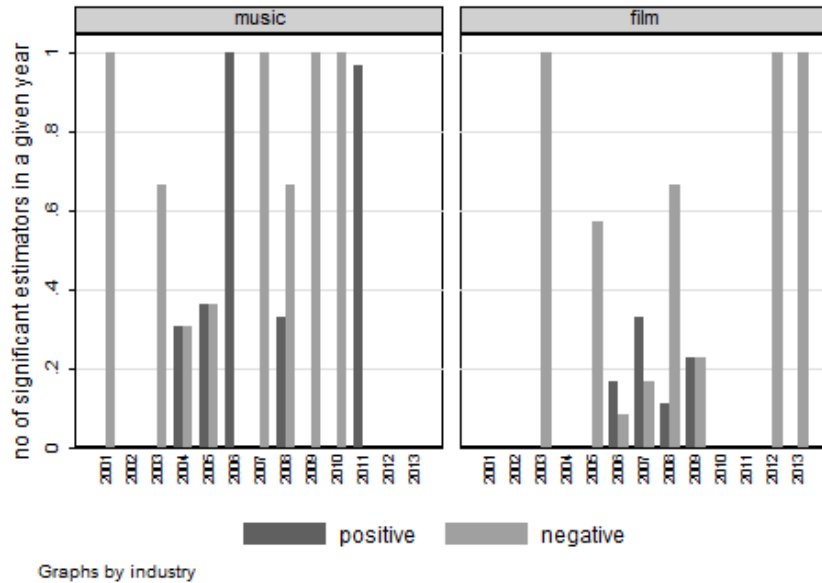Figure 3: Number of significant coefficients, 'reliable' coefficients

Figure 4: Share of significant coefficients, 'reliable' coefficients

Note: Yearly values calculated for all studies, time defined by year encompassed the particular study, un-weighted average. Author's evaluation used to define 'reliable' coefficients to be included. Positive coefficient signifies complementarity between unauthorized acquisition and sales, negative coefficient is consistent with substitution.

This shortage of heterogeneity in the literature may mask substantial cross-country heterogeneity in elasticity of demand in general, that is not adequately accounted for in comparative studies. The country specificity may stem from phenomena related to "piracy" – such as market maturity, quality of domestic cultural goods, knowledge of foreign languages, presence of international commercial authorized service providers, etc – but also from phenomena related to general consumption of cultural goods (such as the average number of hours worked, cultural norms concerning leisure

14

time, etc.).

Similarly, micro-studies are the most frequent for a few countries only, which implies that the estimated effect sizes in our sample may be uninformative for the other countries covered in the literature (they are surely not informative for the countries that were not analyzed by the literature so far).

An important point has been raised recently by Liebowitz (2016) in his overview of the literature on music: the metric of measurement. He argues that if estimates from several studies of music sales are expressed in comparable terms, the effect of file-sharing is able to account for the entire decline of music records sales. The study by Liebowitz (2016) overlaps partially with the studies used in our meta-analysis, but our results are much less definitive about the role of file-sharing. This discrepancy stems from two sources. First, Liebowitz (2016) disregards studies which find a positive effect or no effect of file-sharing on music sales (presenting a thorough justification for such a choice of papers in his overview). Our study is a meta-regression, which by nature necessitates that we cannot disregard any study so long as it provides estimates on a related topic. Second, Liebowitz (2016) focuses on the pre-Napster period, whereas our meta-analysis covers also the studies utilizing data from later periods.

Summarizing, the results do not seem to corroborate consistently any particular side in the debate between the creative industries and the sharing communities. The results of different studies vouch for substitution and complementarity between "online piracy" and legitimate sales, with large dispersion of estimates due to diverse proxies, measures, and identification strategies. An important dimension of heterogeneity may come from the cultural products themselves. The diversification of the cultural goods encompasses challenges such as highly imperfect substitution, preferences instable over time, numerous grounds for complementarity, phenomena such as addiction, etc. A valid example stems from a comparison of two papers covering the case of MegaUpload shutdown. Danaher and Smith (2014) and Peukert et al. (2017) find the opposite effects: the former finds increase in digital downloads (admittedly close substitutes) and the latter finds no change in box office revenue (arguably a less perfect substitute). However, these studies differ also methodologically and in country coverage: 12 for Danaher and Smith (2014) and as much as 50 for Peukert et al. (2017). Given these multiple sources of differences (identification, sales measure and country coverage in this particular example), one may struggle when it comes to formulating ultimate judgments on the nature of the relationship between "online piracy" and legitimate sales.

A related issue concerns awareness: a consumer who paid for a streaming given content (e.g. a tv-series) to an illegitimate distributor may be unaware of illicit character of her consumption, but is likely to change consumption patterns concerning streaming also for other cultural goods (e.g. music or books or film). The growing availability of legitimate streaming sources along with changing culture consumption patterns in general are likely to interact in determining the relationship between "online piracy" and general consumption of cultural goods.

5   Conclusions

There is a number of theoretical insights on how the authorized sales of cultural goods may be affected by the unauthorized distribution of their digital versions online: there may be important complementarities as well as substitution, making the overall effect on authorized sales an empirical question. The representatives of the creative industries race in providing the dollar estimates of the losses due to "online piracy", while the proponents of sharing argue that the decisions to acquire or purchase may be driven by different intrinsic motivations, thus breaking the link between the number of files exchanged without authorization and the sales. The economic profession has taken a number of attempts to tackle this empirical issue. The objective of this study was to summarize these achievements and provide a synthetic, quantitative overview of the overall findings from the studies conducted so far.

Our focus in this study was on the estimates of the unauthorized acquisition on sales. These estimates vary from large positive (signifying complementarity and a plethora of other effects) to large negative (signifying sales displacement). Our systematic overview suggests that there is no clear conclusion so far in the literature. Even if we narrow the analysis to what the authors considered the most reliable set of estimates and consider only peer-reviewed published papers, the results reveal weakly significant negative size effects with strong evidence for publication bias towards negative results, in particular in higher ranked journals. There are also important differences in how technological progress affects the sign and magnitude of the relationship between authorized sales and unauthorized acquisition: studies covering later periods are somewhat more likely to show negative effects on sales in the case of film industry and somewhat more likely to show positive effects in the case of music industry. However, these time effects are not large, given the substantial changes in the business models and identification strategies.

There is a number of reasons why the literature may be so unequivocal after two decades of research. First, both sales and "online piracy" belong to a group of rather poorly measurable phenomena. While in the case of unauthorized acquisition this requires no further explanation, in the case of sales that may come as a surprise, since for most other industries sales or consumption data seem readily available. In the case of the creative industries, however, there is a variety of indicators – such as sales revenues, number of items sold or number of viewers – that are not necessarily similar to one another. They also substitute one another to some extent, which means that the effects on consumption for one distribution channel is affected by the sales effectiveness in other channels, irrespective of the unauthorized acquisition. Further, to make comparisons meaningful, one would ideally want not only data from all channels at once, but also at product level and over longer horizon (to isolate product fixed effects, such as popularity or particular target group). In the literature so far, it is rarely the case. Finally, the industry that possesses the relevant data has a strong interest in proving that the negative impact of piracy is larger than it really is. It is thus incentivized to make the data available in a self-serving manner. It surprising, how few randomized control trials and field experiments have been performed thus far, even though experimenting on the distribution of the newest song or movie seems much less

ethically controversial than e.g. drug testing or policy experiments.

Second, the identification of the causal link remains a challenge. The literature has invoked quasi-natural experiments as the means to overcome this problem, due to the shortage of valid instruments. Yet, in the case of controlled experiments most of the results rely on self-reported purchasing and downloading behavior, whereas in the case of the quasi-natural experiments, authors emphasize that data available suffers from insufficient breadth or quality.

The third – and possibly the most important reason – concerns the underlying intuitions concerning the links between unauthorized acquisition and sales. The phenomenon of "online piracy" seems to be highly complex in nature. Except for the obvious acquisition by the end-consumers, there is also the case of uploading by quasi-professionals, leak piracy, potential piracy and the lag between the piracy and the observed sales. Especially this last point requires more specific treatment in the literature due to the specific characteristics of the cultural goods. On the one hand, these are durable goods in a sense that a song or a movie usually does not need to be repurchased. Yet, if the first-time consumption occurred on television or radio (usually unobservable), consumers may be actually more inclined to purchase their own copy. On the other hand, interest in many cultural goods tends to be short lived, at least for some cultural goods. The arrival rate of this subset of the cultural goods may be compared to the FMCG industry: the song or a movie purchased last week or last month may no longer be of interest for the consumer if a new one arrives (possibly by a different artist and a different producer). Thus, the cultural goods exhibit a joint set of difficulties of durable and non-durable goods, which all affect the consumer decision process and which all exhibit different reaction to the presence of unauthorized, virtually free distribution. Until research is able to disentangle these effects – possibly also by identifying the type of consumers or, better yet, the type of consumption – the results of the studies are at risk of confounding a variety of issues which correlate well with unauthorized consumption, and thus blur the estimates of a causal relationship between "online piracy" and sales.

References

Adermon, A. and Liang, C.-Y.: 2014, Piracy and music sales: The effects of an anti-piracy law, *Journal of Economic Behavior & Organization* **105**, 90–106.

Aguiar, L. and Martens, B.: 2016, Digital music consumption on the internet: evidence from clickstream data, *Information Economics and Policy* **34**, 27–43.

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

Andersen, B. and Frenz, M.: 2010, Dont blame the p2p file-sharers: the impact of free music downloads on the purchase of music CDs in canada, *Journal of Evolutionary Economics* **20**(5), 715–740.

Bae, S. H. and Choi, J. P.: 2006, A model of piracy, *Information Economics and Policy* **18**(3), 303–320.

Bai, J. and Waldfogel, J.: 2012, Movie piracy and sales displacement in two samples of chinese consumers, *Information Economics and Policy* **24**(3–4), 187–196.

Balestrino, A.: 2008, It is a theft but not a crime, *European Journal of Political Economy* **24**(2), 455–469.

Barker, G. and Maloney, T.: 2015, The impact of internet file-sharing on the purchase of music cds in canada, *Journal of Evolutionary Economics* **25**(4), 821–848.

Belleflamme, P.: 2016, The economics of digital goods: A progress report, *Review of the Economic Research on Copyright Issues* **13**(2), 1–24.

Bender, M. T. and Wang, Y.: 2009, The impact of digital piracy on music sales: A cross-country analysis, *International Social Science Review* **84**(3/4), 157–170.

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

Blackburn, D.: 2004, On-line piracy and recorded music sales, *Unpublished manuscript* .

Blackburn, D.: 2006, The heterogeneous effects of copying: The case of recorded music, *Job Market Paper (Harvard Ph. D. Programme)* .

Boorstin, E. S.: 2004, *Music sales in the age of file sharing*, PhD thesis, Princeton University.

Bounie, D., Waelbroeck, P. and Bourreau, M.: 2006, Piracy and the demand for films: Analysis of piracy behavior in french universities, *SSRN eLibrary* .

Chen, Y.-n. and Png, I.: 2003, Information goods pricing and copyright enforcement: Welfare analysis, *Information Systems Research* **14**(1), 107–123.

Chi, W.: 2008, *Does file sharing crowd out copyrighted goods? Evidence from the music recording industry*, November.

Danaher, B., Dhanasobhon, S., Smith, M. D. and Telang, R.: 2010, Converting pirates without cannibalizing purchasers: The impact of digital distribution on physical sales and internet piracy.

Danaher, B. and Smith, M. D.: 2014, Gone in 60 seconds: The impact of the megaupload shutdown on movie sales, *International Journal of Industrial Organization* **33**, 18.

Danaher, B., Smith, M. D. and Telang, R.: 2014, Piracy and copyright enforcement mechanisms, *Innovation Policy and the Economy* **14**(1), 25–61.

Danaher, B., Smith, M. D., Telang, R. and Chen, S.: 2014, The effect of graduated response anti-piracy laws on music sales: evidence from an event study in france, *The Journal of Industrial Economics* **62**(3), 541–553.

De Vany, A. and Walls, W.: 2007, Estimating the effects of movie piracy on box-office revenue, *Review of Industrial Organization* **30**(4), 291–301.

Dejean, S.: 2009, What can we learn from empirical studies about piracy?, *CESifo Economic Studies* **55**(2), 326–352.

Elberse, A.: 2010, Bye-bye bundles: The unbundling of music in digital channels, *Journal of Marketing* **74**(3), 107–123.

Fukugawa, N.: 2011, How serious is piracy in the videogame industry?, *SSRN Scholarly Paper ID 1828642*, Social Science Research Network, Rochester, NY.

Gopal, R. D., Bhattacharjee, S. and Sanders, L.: 2006, Do artists benefit from online music sharing?, *The Journal of Business* **79**(3), 1503–1533.

Grassmuck, V. R.: 2010, Academic studies on the effect of file-sharing on the recorded music industry: A literature review, *SSRN Scholarly Paper ID 1749579*, Social Science Research Network.

Halmenschlager, C. and Waelbroeck, P.: 2012, Fighting free with free: Streaming vs. piracy, *Technical report*, Working paper.

Hammond, R. G.: 2014, Profit leak? pre-release file sharing and the music industry, *Southern Economic Journal* **81**(2), 387408.

Hennig-Thurau, T., Henning, V. and Sattler, H.: 2007, Consumer file sharing of motion pictures, *Journal of Marketing* **71**(4), 1–18.

Higgins, J. P., Green, S. et al.: 2008, *Cochrane handbook for systematic reviews of interventions*, Vol. 5, Wiley Online Library.

Hong, S.-H.: 2013, Measuring the effect of napster on recorded music sales: Difference-in-differences estimates under compositional changes, *Journal of Applied Econometrics* **28**(2), 297324.

King, S. P. and Lampe, R.: 2003, Network externalities, price discrimination and profitable piracy, *Information Economics and Policy* **15**(3), 271–290.

Koh, B., Murthi, B. P. S. and Raghunathan, S.: 2013, Shifting demand: Online music piracy, physical music sales, and digital music sales (shifting demand: Piracy, physical & digital music sales).

Lee, J.: 2009, Private filesharing markets and album sales.

Leung, T. C.: 2015, Music piracy: Bad for record sales but good for the iPod?, *Information Economics and Policy* **31**, 1–12.

Liebowitz, S. J.: 2008, Research note: Testing file sharing's impact on music album sales in cities, *Management Science* **54**(4), 852–859.

Liebowitz, S. J.: 2013, Chapter 23. internet piracy: the estimated impact on sales, *Handbook on the Digital Creative Economy*, Edward Elgar Publishing, p. 262.

Liebowitz, S. J.: 2016, How much of the decline in sound recording sales is due to file-sharing?, *Journal of Cultural Economics* **40**(1), 13–28.

Liebowitz, S. J.: 2017, A replication of four quasi-experiments and three facts from" the effect of file sharing on record sales: an empirical analysis"(journal of political economy, 2007), *Economics: The Open-Access, Open-Assessment E-Journal* **11**(2017-13), 1–22.

Liebowitz, S. J. and Zentner, A.: 2012, Clash of the titans: Does internet use reduce television viewing?, *Review of Economics and Statistics* **94**(1), 234245.

Ma, L., Montgomery, A. L., Singh, P. V. and Smith, M. D.: 2014, An empirical analysis of the impact of pre-release movie piracy on box office revenue, *Information Systems Research* **25**(3), 590–603.

Martikainen, E.: 2014, Does file-sharing reduce DVD sales?, *NETNOMICS: Economic Research and Electronic Networking* **15**(1), 9–31.

McKenzie, J. and Walls, W. D.: 2013, File-sharing and film revenues: An empirical analysis, *Working Paper 2013-14*, University of Sydney, School of Economics.

Michel, N. J.: 2006, The impact of digital file sharing on the music industry: An empirical analysis, *Topics in Economic Analysis & Policy* **6**(1).

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

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

Oberholzer-Gee, F. and Strumpf, K.: 2010, File sharing and copyright, *Innovation Policy and the Economy, Volume 10*, University of Chicago Press, pp. 19–55.

Orme, T.: 2014, The short- and long-term effectiveness of anti-piracy laws and enforcement actions, *Journal of Cultural Economics* **38**(4), 351–368.

Peitz, M. and Waelbroeck, P.: 2004, The effect of internet piracy on CD sales: Cross-section evidence.

Peitz, M. and Waelbroeck, P.: 2006, Why the music industry may gain from free downloadingthe role of sampling, *International Journal of Industrial Organization* **24**(5), 907–913.

Peukert, C., Claussen, J. and Kretschmer, T.: 2017, Piracy and box office movie revenues: Evidence from megaupload, *International Journal of Industrial Organization* **52**, 188–215.

Reavis-Conner, K. and Rumelt, R. P.: 1991, Software piracy: an analysis of protection strategies, *Management Science* **37**(2), 125–139.

Rob, R. and Waldfogel, J.: 2006, Piracy on the high cs: Music downloading, sales displacement, and social welfare in a sample of college students, *Journal of Law and Economics* **49**(1), 29–62.

Rob, R. and Waldfogel, J.: 2007, Piracy on the silver screen, *The Journal of Industrial Economics* **55**(3), 379–395.

Rochelandet, F. and Le Guel, F.: 2005, P2p music sharing networks: why the legal fight against copiers may be inefficient, *Review of Economic Research on Copyright Issues* **2**(2), 6982.

Sandulli, F. D.: 2007, CD music purchase behaviour of P2P users, *Technovation* **27**(6-7), 325–334.

Siwek, S.: 2006, The true cost of motion picture piracy to the u.s. economy, *Policy Report 186*, Institute for Policy Innovation.

Siwek, S. E.: 2007a, *The true cost of Copyright Industry Piracy to the US economy*, Institute for Policy Innovation.

Siwek, S. E.: 2007b, *The true cost of sound recording piracy to the US economy*, Institute for Policy Innovation.

Smith, M. and Telang, R.: 2009, Competing with free: The impact of movie broadcasts on DVD sales and internet piracy, *MIS Quarterly* **33**(2), 321–338.

Smith, M. and Telang, R.: 2012, Assessing the academic literature regarding the impact of media piracy on sales, *SSRN Scholarly Paper ID 2132153*, Social Science Research Network, Rochester, NY.

Stanley, T. D. and Doucouliagos, H.: 2015, Neither fixed nor random: weighted least squares meta-analysis, *Statistics in Medicine* **34**(13), 2116–2127.

Stevans, L. K. and Sessions, D. N.: 2005, An empirical investigation into the effect of music downloading on the consumer expenditure of recorded music: A time series approach, *Journal of Consumer Policy* **28**(3), 311–324.

Tanaka, T.: 2004, Does file sharing reduce music CD sales?: A case of japan, *IIR Working Paper 05-08*, Institute of Innovation Research, Hitotsubashi University.

Tetzlaff, J., Moher, D., Pham, B. and Altman, D.: 2006, Survey of views on including grey literature in systematic reviews, *14th Cochrane Colloquium, Dublin, Ireland.*

Waldfogel, J.: 2009, Lost on the web: Does web distribution stimulate or depress television viewing?, *Information Economics and Policy* **21**(2), 158–168.

Waldfogel, J.: 2010, Music file sharing and sales displacement in the iTunes era, *Information Economics and Policy* **22**(4), 306–314.

Waldfogel, J.: 2011, Bye, bye, miss american pie? the supply of new recorded music since napster, *Working Paper 16882*, National Bureau of Economic Research.

Wang, C.-C., Chen, C.-t., Yang, S.-c. and Farn, C.-k.: 2009, Pirate or buy? the moderating effect of idolatry, *Journal of Business Ethics* **90**(1), 81–93.

Zentner, A.: 2006, Measuring the effect of file sharing on music purchases*, *The Journal of Law and Economics* **49**(1), 63–90.

Zentner, A.: 2008, Online sales, internet use, file sharing, and the decline of retail music specialty stores, *Information Economics and Policy* **20**(3), 288–300.

Appendix A. List of papers

During the course of our search we have encountered a number of studies related to our topic that could not be included in our metaregression. There were also cases of studies that could be used but with some omissions. Those exclusions can be divided into five groups and are based on rules that we followed consistently.

1. Some of the published studies interpret their results as the effect of file-sharing, although later research has raised concerns on whether such interpretation is grounded. On the other hand, some studies with similar specification stated that their results should not be interpreted as evidence in this context. If authors' interpretation is that their study estimates the effect of file-sharing on cultural good sales, we took it at face value to eliminate any room for discretion on which studies are included. However, if another author performed a similar analysis but did not claim that it measures an effect of online "piracy" on sales, then it was not included. For example, Zentner (2008) studies Internet penetration effect on the survivability of music stores. However, this proxy for sales reflects not only the demand but also the supply side. Moreover, the number of stores is a very different type of proxy to the others used in the literature (which all describe some form of direct consumption). Another issue in Zentner (2008) is the interpretation of the piracy proxies (internet penetration AND broadband penetration), which relies heavily on the study design, making it incomparable to other studies).

2. We included the newest version of the study available (if we found a working paper, we would then check if there were any newer – or published – versions of it). If there were additional estimates in the older version – such as Blackburn (2004) compared to Blackburn (2006) – we did not take them into account.

3. We did not include any article that had its results based on non-regression calculations. This led to exclusion of reports prepared for specific institutions and industry representatives (e.g Siwek 2006, 2007a,b).

4. When there were two or more piracy variables in one regression, we would treat each of them as a separate observation. However, in some cases, when the model interpretation was based on e.g. two or more variables *simultaneously* (e.g. regressions with piracy interacted with popularity in Blackburn 2006, or on the comparison between two specifications, it could not be included in our study).

5. We found only one study related to video games effect on sales (Fukugawa 2011). Since it was too little to include it in the analysis, we concentrated on the music and movie industries and left it out of our sample.

Table A.1: List of papers with regression estimates on the link between pirates and sales

Paper	No. of estimates	Relationship
Aguiar and Martens (2016)	36	Positive
Andersen and Frenz (2010)	7	Neutral
Bai and Waldfogel (2012)	28	Varying
Barker and Maloney (2015)	4	Negative
Bender and Wang (2009)	1	Negative
Bhattacharjee et al. (2007)	Not in metaregression	Varying
Blackburn (2006)	3	Varying
Boorstin (2004)	5	Varying
Bounie et al. (2006)	11	Varying
Chi (2008)	29	positive
Danaher et al. (2010)	Not in metaregression	Varying
Danaher, Smith, Telang and Chen (2014)	Not in metaregression	Negative
Danaher and Smith (2014)	12	Negative
De Vany and Walls (2007)	4	Negative
Elberse (2010)	2	Negative
Gopal et al. (2006)	Not in metaregression	Varying
Hammond (2014)	19	Varying
Hennig-Thurau et al. (2007)	18	Negative
Hong (2013)	2	Negative
Koh et al. (2013)	Not in metaregression	Varying
Lee (2009)	1	Positive
Leung (2015)	2	Varying
Adermon and Liang (2014)	16	Negative
Liebowitz (2008)	7	Negative
Liebowitz and Zentner (2012)	13	Varying
Ma et al. (2014)	10	Negative
Martikainen (2014)	2	Neutral
McKenzie and Walls (2013)	25	Negative
Michel (2006)	3	Negative
Mortimer et al. (2012)	Not in metaregression	Varying
Oberholzer-Gee and Strumpf (2007)	42	Neutral
Orme (2014)	2	Neutral
Peitz and Waelbroeck (2004)	32	Negative
Peukert et al. (2017)	11	Varying
Rob and Waldfogel (2006)	14	Negative
Rob and Waldfogel (2007)	30	Negative
Rochelandet and Le Guel (2005)	Not in metaregression	Negative
Smith and Telang (2009)	3	Neutral
Stevans and Sessions (2005)	2	Negative
Tanaka (2004)	8	Neutral
Waldfogel (2009)	54	Negative
Waldfogel (2010)	20	Negative
Wang et al. (2009)	Not in metaregression	Varying
Zentner (2006)	5	Negative
Zentner (2008)	29	Negative

Table A.2: Number of studies and coefficients per year

Year of publication	No. of published articles	No. of working papers	No. of estimates	Conclusions Film	Conclusions Music
2004	0	3	40	0	7
2005	2	0	2	0	2
2006	5	1	36	3	7
2007	5	0	85	5	3
2008	2	1	29	0	3
2009	4	1	41	2	4
2010	4	0	29	2	4
2011	0	2	22	2	0
2012	3	1	32	6	2
2013	1	6	76	5	7
2014	3	1	33	2	15
Total	29	16	425	27	54

Note: whenever a published paper is available, the working paper is disregarded.

Table A.3: Measures and proxies

Proxy for sales (as dependent variable)		No. of papers	No. of estimates
Consumer side	Viewings	6	107
	Purchases	11	98
	Clicks on authorized websites	1	36
	Expenditure	3	7
Producer side	Sales	8	109
	Revenues	8	65
	Rank	1	3

Proxy for piracy (as independent variables)		No. of papers	No. of estimates
	Downloads	18	258
	Spread of piracy	4	39
	Clicks on unauthorized websites	1	36
	Tech/Law change	6	40
	Internet/tech proficiency	1	29
	Supply	5	23

Note: Tech/Law change refers to studies of impact of a huge dimension in the market for pirated products, these include impositions of new laws and shutdowns of big piracy networks like Megaupload or Napster. Supply signifies studies which used measures such as owning equipment (e.g. CD burners, P2P software). In Table A.4 we tabulate sales proxies against the piracy proxes. This tabulation demonstrates that only a small number of combinations of these measures actually exists in the literature.

Table A.4: The combinations of sales and "piracy" measures in the literature

Proxy for the sales variable	Proxy for the piracy variable							
	Clicks	Downloads	Know-how	Legislative change	Other	Spread of piracy	TLS change	Total
Clicks (on authorized websites)	36	0	0	0	0	0	0	36
Expenditure (declared money spent)	0	2	0	0	0	0	5	7
Purchases (declared units bought)	0	65	29	0	0	4	0	98
Rank (chart or sale)	0	0	0	0	3	0	0	3
Revenues (money earned)	0	25	0	10	14	3	13	65
Sales (units sold)	0	68	0	0	3	32	6	109
Watching (one-time consumption)	0	98	0	0	3	0	6	107
Total	36	258	29	10	23	39	30	425

Notes: see note under Table A.3.

Appendix B. Regressions full set of estimates

Table B.5: Regressions for the film industry

	RE						WLS					
	Film			Music			Film			Music		
	ALL	ALL with IF		ALL	ALL with IF		ALL	ALL with IF		ALL	ALL with IF	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Constant	-10.502	-64.569	-157.895***	-16.552	-41.037	-158.267***						
	(26.241)	(50.141)	(56.058)	(16.067)	(29.702)	(23.803)						
Precision							34.748***	17.711	-5.897	-16.063***	12.785*	-69.576**
							(10.219)	(21.308)	(24.366)	(5.266)	(7.110)	(34.878)
Journal's impact factor			-0.564***			0.155***			-0.081*			0.074**
			(0.191)			(0.023)			(0.042)			(0.031)
Period of study (first year)	-0.011	0.022	0.189***	-0.017**	-0.094***	-0.118***	-0.010	-0.003	0.027	0.004***	0.063***	0.027
	(0.011)	(0.025)	(0.061)	(0.008)	(0.029)	(0.018)	(0.012)	(0.009)	(0.018)	(0.002)	(0.014)	(0.020)
Year of publication	0.017	0.011	-0.107**	0.026**	0.114***	0.196***	-0.006	-0.005	-0.022**	0.004	-0.069***	0.008
	(0.013)	(0.014)	(0.042)	(0.012)	(0.039)	(0.028)	(0.013)	(0.005)	(0.011)	(0.004)	(0.016)	(0.036)
Study span (in years)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of variables used	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of the observations	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample selection acknowl.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Reverse causality acknowl.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE for sales proxies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cinema based study	Yes	Yes	Yes			Yes	Yes	Yes				
FE for "piracy" proxies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# of coefficients	179	131	131	242	125	125	178	130	130	221	123	123
R-squared							0.236	0.192	0.216	0.412	0.300	0.334

Notes: Columns (1) to (6) obtained from `-metareg-` in STATA. Columns (7)-(12) show weighted least squares following Stanley and Doucouliagos (2015), with LHS and RHS variables divided by the standard error of the effect. Standard errors in parentheses. *** $p<0.05$, ** $p<0.10$, * $p<0.15$.