

THE EFFECT OF INTERNET PIRACY ON MUSIC SALES: CROSS-SECTION EVIDENCE

MARTIN PEITZ AND PATRICK WAELEBROECK

ABSTRACT. We use a 1998-2002 cross-section dataset to analyze the claim of losses due to internet piracy made by the record industry. The results suggest that internet piracy played a significant role in the decline in music sales during the early days of file-sharing networks.

1. INTRODUCTION

Many voices in the music industry have claimed that internet piracy has reduced sales of legitimate CDs and that illegal MP3 downloads have become a substitute for legal CD purchases.¹ Indeed, many analysts believe that the current downturn in CDs sales is due to the increasing and uncontrollable number of illegal copies available using peer-to-peer (P2P) technologies. End-user piracy, which is different from commercial piracy, seems to be much more difficult to control. Representatives of the music industry have claimed that the very existence of the industry is at stake.² Policy-makers and record companies have addressed internet piracy by reinforcing copyright laws, by implementing technological protection and by actively enforcing legal protection through lawsuits targeted at developers and users of P2P networks. Yet, defenders of online distribution technologies have argued that MP3 downloads offer a new way for consumers to try new music. By exposing internet users to a larger variety of products, file-sharing technologies could allow them to increase the set of albums they have a potential interest in, thus increasing music sales. This is referred to as sampling or the exposure effect in the economics literature.³ It remains to determine the net contribution of MP3 downloads on music, which is the purpose of this article.

To our knowledge, there have been only two parallel attempts to assess the role of online piracy on music sales. Zentner (2003) uses individual survey data from October 2001 in major European countries. After controlling for unobserved heterogeneity in music taste, he finds that music downloads reduce the probability

¹In the news-breaking Napster case, the court found that the use of the file-sharing technology harmed the music industry partly on the ground that illegal music download reduced legitimate sales of CDs. During the summer 2003, the Recording Industry Association of America (RIAA) started to monitor activities of file-sharing networks in preparation of a massive series of lawsuits targeting individual file-swappers. Eventually, in September 2003, the RIAA initially filed 261 lawsuits charging music uploaders with copyright infringement. By October 2004, the total number of such lawsuits has reached a total of 6191. These legal actions are based on the belief that music downloads are causing *substantial* damage to the music industry.

²Liebowitz (2003) discusses this annihilation hypothesis.

³Peitz and Waelbroeck (2003) offer a critical review of the literature on piracy of digital products. For a particular model that focuses on the sampling aspect, see Peitz and Waelbroeck (2004a).

of purchasing music by 30%. Assuming that people who download music tend to purchase as much as people who do not, Zentner finds that internet piracy could have decreased CD sales by 7% in the countries considered. However, Zentner's dataset only measures music purchases as a binary variable and does not capture the intensity of music downloads.

Oberholzer and Strumpf (2004) use actual download and sales data. They determine which albums have been downloaded most on file-sharing networks during the last quarter of 2002. Controlling for possible endogeneity issues, they show, contrary to the previous study, that the number of times an album has been downloaded does not have an economically significant effect on sales. This study has been criticized by some academics and representatives of the music industry. Liebowitz (2004) argues that the effect of file-sharing on sales of individual albums is hard to extrapolate at the industry level and questions the validity of the instruments chosen by the authors. IFPI market research director Keith Jopling quoted by BBC News ("Legal song downloads rise tenfold", April 1, 2004) criticizes the choice of the last quarter of the year to carry out an empirical study because of the changing nature of music sales due to Christmas.

In this article, we assess the effect of internet piracy on music sales in two steps after describing the data and the methodology in Section 2. First, in section 3 we find a significant negative effect of MP3 downloads on music sales during the period 1998-2002. Second, we discuss the effect MP3 downloads on music sales worldwide as well as for the US in section 4. Section 5 concludes the article.

2. CROSS-SECTION ANALYSIS

2.1. Factors influencing cross-section variation in CD sales. There are several factors in addition to internet piracy that could influence cross-country variation in music sales:

1. prices;
2. country-specific environment of the music industry, including offline commercial piracy, taste for music, the distribution of income of potential users;
3. income and economic environment;
4. substitution with other media and other forms of entertainment; new distribution channels; new media such as DVDs.
5. the "quality" of music.

Liebowitz (2003) reviews these factors for the US. In Peitz and Waelbroeck (2004b), we find that the real price of music has not followed any significant trend over the recent years, so that factor (1) will not play a significant role. Price differences across countries and other country-specific factors (2) that have not changed over the period will be eliminated by taking first differences. Factor (3) will be controlled by domestic GDP. Factors in (4) will be partly captured by the household penetration of DVD and of Digital Music Players. Factor (5) is unobservable and will be assumed to be uncorrelated with the internet piracy variable.

2.2. Data and variable description. We use data on the music industry from the IFPI World Report of 2003 that provides industry data from 1998 to 2002. It does not seem relevant to look at data prior to 1998 since file-sharing technologies, a prerequisite for large-scale internet piracy, only appeared in the second half of 1999 with the creation of Napster. Numbers on music downloads have been obtained from IPSOS-REID. The 16 countries used in the empirical study represent countries

with the largest markets for pre-recorded music (in value), accounting for more than 90% of the world market value.⁴

The dependent variable is *Music Sales*, given in total units of pre-recorded music (Singles, LPs, Music cassettes, CDs). Our explanatory variables are summarized in Appendix 1:

- *GDP* (in constant dollars) captures the economic environment.⁵
- *Downloads*, defined as the percentage of adult internet users who downloaded music files in MP3 format from the Internet at least once.⁶ This variable serves as a proxy for online end-user piracy but does not completely capture the intensity of the downloading activity: ideally, this number should be multiplied by the number of downloaded music files per user in each country. Implicitly, we assume that the effect of an additional downloader on CD sales is similar across the countries in our sample.
- *Broadband*, defined as the number of people with home broadband connection in percentage of the total number of households. Broadband penetration partly captures the substitution with new forms of online activities, such online games as well as audio-streaming and legal downloads from commercial internet sites.⁷ It may also partly capture the intensity of the downloading activity because broadband makes music downloads less time-consuming.
- *Digital media players* (DMP), defined as the percentage of households with a device to play digital music files, such as MP3 players, disk-based portable players or Net MD players. Digital music players account for the transitional effect of new devices that enhance the value of music downloaded from the internet.⁸

⁴These countries are US, Japan, UK, Germany, France, Canada, Mexico, Spain, Italy, Australia, Netherlands, Sweden, South Korea, Belgium, Switzerland, and Taiwan. While we have missing observations for Austria ranked 15th in market value, Brazil (ranked 7th with 1.7% of worldwide sales) was an outlier with almost no internet penetration that we discarded from the econometric analysis. Nevertheless, the partial effect of internet piracy was identical with or without Brazil, although the fit was much poorer by including Brazil in the analysis. There are several explanations for Brazil being an outlier. First, Brazilian GDP is a poor measure of disposable income across the population. Also, commercial piracy is prevalent in Brazil and would need to be taken into account. Finally, Brazil has a very low internet penetration rate with virtually no broadband penetration.

⁵GDP data are from the OECD except for Taiwan for which we gathered compatible information from the Economist Intelligence Service.

⁶These are survey data from the beginning of 2002 for all the countries included in our sample (documented in IPSOS-REID Tempo, 2002). Only the share of downloaders among internet users is reported. This is a unique dataset which gives us a meaningful proxy for online piracy. Unfortunately, this restricts our analysis to a cross-section analysis. Music downloaders defined as the percentage of internet users captures better the volume of internet piracy than when it is defined as the percentage of the total population. See our discussion in the estimation results.

⁷We would like to stress that before 2003, legal downloads were almost inexistent. Online music stores backed by major technology and retail companies only appeared in the second quarter of 2003.

⁸Data were missing for South Korea, Taiwan and Mexico. For Mexico, we used the percentage from Spain (3%), for Korea and Taiwan, we respectively used 19% and 17%, which are the second and third highest penetration rate after Japan (68%).

- *DVD players*, defined as the percentage of households with a DVD player. DVD penetration captures substitutions with new forms of home entertainment.⁹
- *CD-R* is defined as the percentage of households with CD-R devices. CD-R penetration captures the combined effect of online and offline piracy due to people burning MP3 downloaded from the internet and giving them to friends as well as the extent of copying original CDs or CD-Rs from friends.

2.3. Model. We have $n = 16$ countries during the period 1998-2002. Let $y_{i,t}$ denote music sales in log units in country i , $i = 1, \dots, n$, at time t . $z_{i,t}$ is the variable that captures internet piracy (in logs). $x_{i,t}$ is a vector of dimension $k \times 1$ of control variables (also in log) in country i at time t . To eliminate country specific effects, we write the model in first difference:

$$y_{i,t} - y_{i,t-1} = (x_{i,t} - x_{i,t-1})' \beta + \alpha(z_{i,t} - z_{i,t-1}) + \varepsilon_{i,t}$$

where $\varepsilon_{i,t}$ is an unobservable variable with mean 0 and variance σ^2 that includes factor (5) assumed to be uncorrelated with our explanatory variables, and α and β are parameters to be estimated. More specifically, α is the partial effect of online piracy on music sales.

We do not have data on MP3 downloads, Broadband, DMP and DVD variables for the years prior to 2002. Thus we used these variables in level in the regressions. However, levels are likely to be a good proxy for first differences in that period for most countries: in particular, music downloading on file-sharing networks and broadband access essentially started in 1999.

3. ESTIMATION RESULTS

Table 1 gives our estimation results (ordinary least squares with the same weight for each country). We present the result without a constant intercept. With a model in differences, this constant term corresponds to a time trend of the factors not included in our regression. We feel confident that we have accounted for most of the factors that could potentially influence music sales.¹⁰

⁹All DVD data were available for 2003, but some were missing for 2002. We reconstructed these missing data using the growth rate of equivalent countries for 2002-2003. Nevertheless the results were almost identical whether we used the reconstructed DVD variable for 2002 or the observed DVD penetration in 2003.

¹⁰We also ran regressions with a constant term. Results were unreliable as the constant term was estimated to be between -10 and -50 and was most of the time non-significantly different from zero. Moreover, some coefficients changed from a negative to a positive sign. Data are available upon request if the reader wants to test additional specifications.

		GDP	Downloads	Broadband	DVD	DMP	CD-R	R^2	Losses
(1)	Coef.	2.03*	-0.53*	-1.65				0.72	-19.81
(1)	t-stat.	3.30	-2.63	-3.41					
(2)	Coef.	1.99*	-0.39	-1.64*	-0.26			0.74	-14.83
(2)	t-stat.	3.21	-1.59	-3.38	-0.92				
(3)	Coef.	1.91*	-0.45*	-1.47*		-0.27		0.75	-17.02
(3)	t-stat.	3.15	-2.22	-2.99		-1.30			
(4)	Coef.	2.04*	-0.45	-1.61			-0.15	0.72	-16.84
(4)	t-stat.	3.19	-1.44	-3.15			-0.34		
(5)	Coef.	1.90*	-0.37	-1.49	-0.19	-0.24		0.76	-13.81
(5)	t-stat.	3.05	-1.49	-2.94	-0.65	-1.08			
(6)	Coef.	2.00*	-0.39	-1.64*	-0.26		-0.01	0.74	-14.68
(6)	t-stat.	3.07	-1.21	-3.15	-0.82		-0.02		
(7)	Coef.	1.88*	-0.54	-1.48		-0.32	0.19	0.76	-20.33
(7)	t-stat.	2.97	-1.73	-2.90		-1.26	0.38		
(8)	Coef.	1.85*	-0.48	-1.51	-0.23	-0.31	0.31	0.77	-18.22
(8)	t-stat.	2.85	-1.48	-2.89	-0.76	-1.19	0.57		

*=significant at the 10% level.

We find a strong positive effect of income (GDP growth) on CD purchases, although this variable may capture other factors related to the domestic economic environment. The coefficient associated with the internet piracy variable has a significant and negative effect on music sales at the 10% confidence level. Based on the estimated coefficient associated with the *Downloads* variable, we can assess the effect of internet piracy on the decline in music sales (last column of Table 1). The implied loss of CD sales due to MP3 downloads is -20% for the period 1998-2002 with specification (1).¹¹ *Broadband* penetration is a significant explanatory variable of music sales and always has a negative impact. We interpret this result as follows. Although broadband penetration is clearly correlated with the intensity of file-sharing activities, we believe that the main effect of internet piracy is captured by our *Downloads* variable. Other factors are related to broadband penetration. Indeed, people are already listening more to audio clip and internet radio than downloading music files. While it is not clear how audio streaming will affect record companies in the future (copyright owners collect royalties on streamed content), it is only one of the many activities that broadband users are doing on any given day. Other forms of digital activities include creating online content (pictures, web pages), watching video clips and movies, playing online games (especially popular among teenagers and college students), purchasing products online and undirected browsing. These new forms of entertainment that have been embraced by broadband users are clearly a substitute to traditional forms of entertainment. In Peitz and Waelbroeck (2004b) we document how intensive internet users have already reduced the amount of time watching television and listening to pre-recorded music. *Digital media players* also negatively influence music sales. Thus there could be a continued effect of music downloads on music sales due to people purchasing new equipment to listen to digital music, until the diffusion of DMP is completed. *DVD*

¹¹To obtain this percentage, we multiplied the estimated coefficients by the sample average number of downloaders as a percentage of the number of internet users (37.69).

penetration does not increase the R^2 that is fairly high given our small number of explanatory variables and the fact that the model is taken in first differences. Finally, using $CD-R$ as a combined measure of online and offline end-user piracy does not add much to the explanatory power of the model: the R^2 basically remains constant. We did not introduce measures for commercial piracy as it seems to be of little relevance for most countries in our sample.

We ran several robustness checks. In particular, we ran regressions using music downloads defined as a percentage of total population, i.e. by multiplying the number of people downloading music as a percentage of internet users by the percentage of households with home internet connection. This variable was not significantly different from zero at the 10% level and had an associated coefficient that sometimes was estimated to be positive when additional variables were included. We argue that this variable does not identify the volume of internet piracy as well as the variable defined in percentage of internet users. Our argument is the following. The number of internet users as a share of the population is based on the number of households with internet connection. This is however misleading since many people do not access the internet from home. Indeed, until 2002 many people were downloading music files from work or school, from an internet café or from a friend's home. While our implicit assumption that internet users as a share of total population is constant across countries is certainly heroic, the alternative to use the number of internet connections from home to calculate the number of internet users as a share of the population does not seem to be a good approximation of the number of people who could effectively access the internet prior to 2002.

4. DISCUSSION

We now discuss the effect of internet music piracy on the evolution of music sales in the U.S. for 2003.

	Kazaa	WinMX	iMesh	Bear-share	Grokster	eMule	Bit-torent	Total	% change
Dec. 2002	32000	7500	1600	1300	1050	50	0	43500	
Jun. 2003	35000	6250	1400	1250	550	150	0	44600	2.5
Dec. 2003	24750	5800	800	1200	400	200	200	33150	-25.7

Source: comScore/Media Matrix; figures in thousands of applications actively running during a given month.

Table 2 reports the evolution of the number internet users with file-sharing application actively running during a given month. We want to relate this evolution to the number of CD sales (and music sales in other formats). The problem in doing so is that file-sharing usage numbers are snapshots in time whereas music sales refer to longer periods (a semester or a full year). This is complicated by the fact that music sales are markedly different in the first half of a year compared to the second half.¹² Using numbers from the IFPI Report “The Recording Industry in Numbers 2004”, total units dropped from 2002 to 2003 by 7.6% (from 849 million units to 789.5 million units). Since this is a percentage change for the full year, it is perhaps more meaningful to compare the sales in the second half of 2002 to the

¹²Note that the second semester is usually stronger than the first semester because of back to school offers, thanksgiving and Christmas events.

sales in the second half of 2003. Here total unit sales hardly changed (457.1 million units in the second semester of 2002 and 453.3 million units sold in the second half of 2003). Hence, the claim of the music industry that illegal music downloads are hurting music sales is confirmed to the extent that in the previous period in which file-sharing was increasing, sales suffered a lot, while in the subsequent period in which usage of file-sharing applications was lower, sales did not decrease significantly. Clearly, this ignores any other factors. Moreover, this exercise can only give an indication of the likely impact of file-sharing. More micro-studies are needed.¹³

With respect to world sales, we used the number of infringing music files on P2P networks tracked by the IFPI as a measure of “internet music piracy” (see Table 3). A picture similar to the situation in the U.S. emerges. At the end of 2003, the number of available infringing music files dropped to the level reached at the end of 2002. Again this evolution may partly explain the fact that total unit sales were flat in the second half of 2003 (1727 million units) compared to the second half of 2002 (1757 million units), while sales were falling in previous periods.

June 2004	700
January 2004	800
June 2003	1000
April 2003	1100
November 2002	900
June 2002	500
April 2002	600

Source: IFPI, World News, 2004; figures in millions of units available at any time.

Other dimensions may have to be considered when assessing the impact of internet piracy on music sales using variables that do not capture the *intensity* of downloading. In particular, there has been self-selection. Teenagers and college students with low purchasing power have the highest propensity to use file-sharing technologies and for this reason adopted the technology first. Older internet users are late adopters with higher purchasing power and high opportunity cost of using file-sharing networks to download music. This could explain why the effect of music downloads has been strongest in the beginning of the file-sharing era. This interpretation is compatible with the study of Boorstin (2004) who finds that the number of teenagers and adults younger than 24 who have internet access significantly decreases total CD sales in a given area, but that the total number of older adults with internet access significantly increases total CD purchases. However, with respect to the study by Boorstin, it may be problematic to equate internet access to internet piracy, as we have argued that internet access can serve a number of purposes, only one of them is downloading music from file-sharing networks. Moreover, analyzing the effect of internet access of a subpopulation on total sales does not provide the correct partial effect of that sub-population.

¹³Contrary to the numbers presented in Table 2, Karagiannis et al. (2004) found that the number of internet users running file-sharing applications has mainly remained constant over the last couple of years. This contradictory result stems from the increased difficulty to accurately measure P2P traffic of the newest generations of file-sharing technologies such as BitTorrent and eDonkey, which is more a concern for 2003 onwards than prior to 2003. Also, the study does not provide evidence on the relative importance of music files vs. films in DIVX format exchanged on P2P sites.

From a more global perspective, the music industry has experienced several technological cycles related to the substitution between different music formats. Cassettes have replaced LPs. CDs have replaced cassettes. Moreover, consumers have for a long time replaced their music collections owned on the LP format by purchasing the same albums on CDs. It seems that this substitution pattern is approaching an end. New formats, such as compressed digital music protected by Digital Rights Management, super-audio CD and music on DVD, have not (yet) taken off.

5. CONCLUSION

We have analyzed the RIAA's claim that music downloads are causing a substantial decrease in music sales. Our macro data confirm their fear: we find that music downloading could have caused a 20% reduction in music sales worldwide between 1998-2002. While this is only a crude estimate, we believe that it is a good reference value that other studies, especially microeconomic ones, could use to assess the exact substitution that has taken place between CDs and MP3s. Our analysis also reveals that other factors than music downloads on file-sharing networks are likely to be responsible for the decline in music sales in 2003.

REFERENCES

- Boorstin, E.** (2004), "Music Sales in the Age of File Sharing", Senior thesis, Princeton University.
- Karagiannis, T., A. Broidon, N. Brownlee, K. Claffy and M. Faloutsos,** (2004), "Is P2P Dying or just Hiding?", Mimeo, University of California, Riverside.
- Liebowitz, S.** (2003), "Will MP3 Downloads Annihilate the Record Industry? The Evidence so Far", in G. Libecap (ed.), *Advances in the Study of Entrepreneurship, Innovation, and Economic Growth*, JAI Press; pp.229-60..
- Liebowitz, S.** (2004), "Pitfalls in Measuring the Impact of File-sharing", Mimeo, University of Texas at Dallas.
- Oberholzer, F. and K. Strumpf** (2004), "The Effect of File sharing on Record Sales: An Empirical Analysis", Mimeo, University of North Carolina at Chapel Hill.
- Peitz, M. and P. Waelbroeck** (2003), "Piracy of Digital Products: A Critical Review of the Economics Literature", CESifo Working Paper #1071.
- Peitz, M. and P. Waelbroeck** (2004a), "Making Use of File Sharing in Music Distribution", Mimeo, University of Mannheim.
- Peitz, M. and P. Waelbroeck** (2004b), "An Economist's Guide to Digital Music", CESifo Working Paper #1333.
- Peitz, M. and P. Waelbroeck** (2004c), "The Effect of Internet Piracy on Music Sales: Cross-Section Evidence", CESifo Working Paper #1122.
- Zentner, A.** (2003), "Measuring the Effect of Online Piracy on Music Sales", Mimeo, University of Chicago.

Appendix. Data

The dataset is summarized in Tables A1 and A2. CD sales and GDP are in log-differences times 100 (i.e. they give percentage change), Downloads 1 and 2, Broadband, DVD, DMP and CD -R are in percentage. Downloads 1 is the variable defined in Section 2, i.e. the number of music downloaders as a percentage of the internet population. Downloads 2 is the number of downloaders as percentage of

total population that we obtained by multiplying Downloads 1 by the percentage of households with home Internet access.

variable	mean	Std.	Min	Max
Music sales 1998-2002	-13.38	19.25	-56.93	15.12
GDP 1998-2002	11.00	6.76	2.98	31.93
Downloads 1	17.90	7.32	5.20	28.00
Downloads 2	37.69	8.13	20.00	51.00
CD-R 2002	23.14	7.72	7.13	39.00
Broadband 2002	9.09	8.95	1.20	32.00
DVD 2002	18.76	11.86	3.01	47.27
DMP 2002	14.06	15.58	2.00	68.00

Sources: IFPI World Reports 2002-2004 (Music sales, Broadband, CD-R, DVD, DMP), OECD (GDP), ISPOS-REID (Downloads).

	Downloads 1	Downloads 2	Broadband	CD-R 2002
Downloads 1	1.00			
Downloads 2	0.69	1.00		
Broadband	0.62	0.67	1.00	
CD-R 2002	0.47	0.18	0.53	1.00

Sources: IFPI World Reports 2002-2004 (Broadband, CD-R), ISPOS-REID (downloads).

MARTIN PEITZ; INTERNATIONAL UNIVERSITY IN GERMANY, SCHOOL OF BUSINESS ADMINISTRATION, 76646 BRUCHSAL, GERMANY, MARTIN.PEITZ@I-U.DE. PATRICK WAELEBROECK, ECARES, UNIVERSITÉ LIBRE DE BRUXELLES, CP 114, 50 AV. ROOSEVELT, 1050 BRUSSELS, BELGIUM, PWAELEBRO@ULB.AC.BE. WE WOULD LIKE TO THANK AN ANONYMOUS REFEREE FOR HELPFUL COMMENTS. THIS PAPER REPLACES AN EARLIER VERSION IN WHICH WE FOCUSED ON THE PERIOD 2001-2002 (SEE PEITZ AND WAELEBROECK, 2004C).