

CD sales and Internet Piracy: Cross-section Evidence

Martin Peitz^{*}

University of Mannheim

Patrick Waelbroeck^{**}

ECARES, Free University of Brussels

June 2003

First draft, do not cite

Abstract: We analyze the role of music downloading in the current downturn in CD sales. After gathering facts on internet piracy in the U.S.A., we provide cross-country evidence of its role on CD sales. We find that music downloading can be associated with a 1.76% decrease in CD sales worldwide between 2000 and 2001 and that Mp3 can explain 25% of the CD decline in the top 10 markets for recorded music.

Keywords: piracy, music, peer-to-peer, internet, survey data, cross-country regressions

JEL-Classification: L

^{*} Address: Department of Economics, University of Mannheim, 68131 Mannheim, Germany,
peitz@bigfoot.de

1 Introduction

Many voices of the music industry [see for instance IFPI, RIAA, IPSOS-REID reports] have claimed that internet piracy has reduced sales of legitimate CDs and that illegal Mp3 have become a substitute to legal CD purchases [see Liebowitz (2002)]. Many analysts believe that the current downturn in CDs sales is due to the increasing and uncontrollable number of illegal copies available using P2P technologies. This belief is illustrated by the current reinforcement of the copyright law, the implementation of technological protection as well as the active enforcement of legal protection through lawsuits targeted at the p2p technologies as well as their users. Yet, defenders of online distribution technologies have argued that mp3 technologies offer a new way for consumers to try new products and to make more informed purchases, which should eventually boost CD sales. This is referred to in the economics literature as sampling or exposure effect.¹

To our knowledge, there has been only two attempts to assess the role of piracy on CD sales. Hui et al. (2002) test the role of *offline* piracy on the demand for CDs. They first build a theoretical model of piracy that includes heterogeneous valuation of consumers, expected fine and quality differential between the original and the copy, as well as a reduced form for the positive influences of illegal copies on the demand for the original, such as network externalities. Secondly, they test whether the demand for originals is increasing or decreasing with the extent of piracy, using international data from the IFPI (28 countries, 1994-1998), the BSA (Business Software Alliance) and the SIIA (Software and Information Industry Association) to construct measures of (offline) piracy and expected fine per country. They find a negative and significant effect of piracy on legitimate CD sales. The implied demand loss is found to be 42% less than the IFPI number and the total revenue loss is evaluated to be no more than 6.6%.

Liebowitz (2002) is the only article that has directly addressed the issue of the effect of *online* piracy on the music industry. He looks at a 30-year time series of sales in the U.S.

** Address: ECARES, Université Libre de Bruxelles, CP 114, 50 av. Roosevelt, 1050 Brussels, Belgium, pwaelbro@ulb.ac.be

¹ See Peitz and Waelbroeck (2003a) for a critical review of the literature on piracy of digital products.

record industry using RIAA numbers until 2001. He argues that the current downturn in CD sales can be associated with the appearance of file-sharing technologies, although he admits that sales of other media much less substitutable to MP3 such as cassettes and singles also dropped during that period. He also dismisses the role of income on CD purchases. He concludes, *without using direct information on downloads or internet usage*, that mp3 can be held responsible for the current decline in record sales and that a further drop of 20% due to mp3 downloads should be observed in the near future.

In this article, we propose to test the effect of internet piracy on CD sales. We proceed in two steps. In section 2, we use cross-section data available since the emergence of P2P technologies in 15 countries. We interpret our results for the US where we have detailed factual survey data on music downloading in the U.S in section 3. Section 4 concludes the article.

2 Cross-section evidence

2.1 Data

We use data from the IFPI World report from 2001 and 2002 that reports data for 2000 and 2001 respectively. It does not seem relevant to look at data much prior to 2000 since P2P technologies only appeared in the second half of 1999. The 15 countries used in the empirical study represent countries with the largest markets for recorded music (in value), accounting for more than 90% of the world total CD market value.² These countries are OECD countries with the exception of Taiwan. We use the number of units sold instead of the value of sales for two reasons. First, sales value is calculated using list prices and not actual purchase prices. Secondly, the real price of a CD has been relatively constant over the last couple of years across different countries (see Liebowitz, 2002).

² While we have missing observations for Austria and Switzerland respectively ranked 15 and 17 in market value, Brazil (ranked 7 with 1.7% of worldwide sales) was an outlier that we discarded from the econometric analysis. Internet piracy is unlikely to be a factor explaining CD sales in Brazil as the internet and broadband penetration are respectively below 2.8% and 0.2% respectively.

We used the following variables in the regressions (Table 1).

CD sales (in units, $\log(\text{CD sales}+1)\times 100$): dependent variable (first difference);

Internet penetration (in percentage of the total population): captures the penetration of mp3 of users of p2p who only download few songs;

Broadband penetration (in percentage of the total number of households): captures heavy users of p2p;

MC sales (in units, $\log(\text{MC sales}+1)\times 100$): to capture cycle effects and replacements of MCs by CDs (first difference);

CD player (numbers of per household): to capture CD market saturation.

We completed information from the IFPI with GDP data from the OECD and the Economist Intelligence Service (for Taiwan).

GDP (in constant dollars, $\log(\text{GDP}+1)\times 100$): captures the domestic economic environment (first difference);

Finally, we provide an additional measure of internet piracy. Data comes this time from an IPSOS-REID Survey done in 2002.

Downloads (18-25, 18+): Represents the percentage of internet users who downloaded music from the Internet at least once.

Table 1. Data ($n = 15$)

Country	GDP	Internet	Broadband	MC sales	CD players	Downloads 18-25	Downloads 18+	CD sales
USA	0.25	62	9.9	-52.40	259	73	40	-6.65
Japan	-0.58	26	7.3	-6.34	123	34	20	-11.20
UK	1.94	38	1.2	-90.63	148	48	28	8.10
Germany	0.57	42	4.5	9.44	125	49	34	-10.73
France	1.82	37	2.6	-25.75	125	57	34	10.78
Canada	1.45	60	20.5	-84.35	113	76	44	-7.76
Mexico	-0.27	18	1.9	-70.00	37	49	35	-7.32
Spain	2.63	29	3.4	-33.61	88	51	39	5.49
Italy	1.76	38	2.9	-32.93	56	70	46	-2.85
Australia	3.86	53	6	-10.43	81	61	37	12.45
Netherlands	1.26	61	9.5	-38.95	126	67	40	-7.93

Sweden	0.82	70	7.4	-27.96	103	75	40	0.76
South Korea	2.99	52	32	-37.37	74	74	50	16.66
Belgium	0.75	44	7.6	0.00	105	55	37	-7.44
Taiwan	-1.91	46	25	-222.3	104	76	51	-22.87
Average	1.16	45.07	9.45	-48.24	111.13	61.00	38.33	-2.03
Std. dev.	1.47	14.70	9.15	56.25	50.35	13.13	7.98	10.79

The last two lines of Table 1 give descriptive statistics of our variables. Note the decrease in CD sales and the much larger drop in MC sales during the period. We do not have data on the Internet, Broadband, CD player and Downloads variables for the year 2000. Thus we use these variables in level in the regressions. The implicit assumption is that the level is a proxy for the change in differences. Note also that we will use music downloads as such and weighted by the population of internet users, leaving the data to tell use which variable provides the best statistical fit.

Table 2 gives a measure of the coherence between our different measures of online piracy.

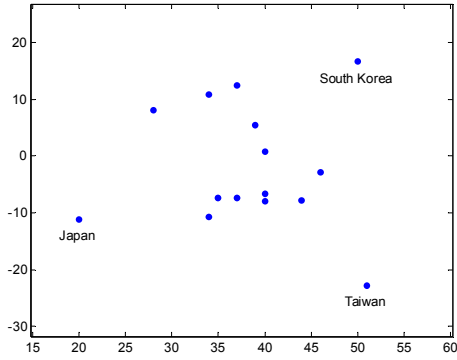
Table 2. Correlation between online piracy variables ($n = 15$)

	Internet	Broadband	Downloads 18-25	Downloads 18+
Internet	1.0000	0.4745	0.7298	0.2985
Broadband	0.4745	1.0000	0.6293	0.6033
Download 18-25	0.7298	0.6293	1.0000	0.8278
Download 18+	0.2985	0.6033	0.8278	1.0000

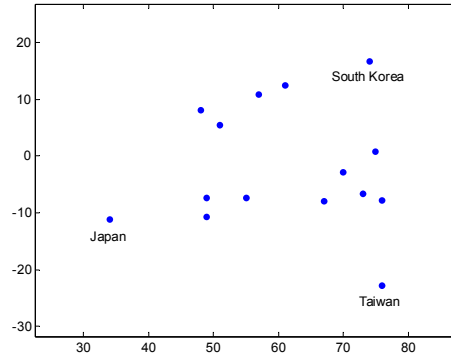
Table 3 displays the partial correlation between the dependent variables on our measures of online piracy. On the one hand, there is strong positive linear correlation between CD sales and economic growth and a negative relation between CD sales and music downloads (among internet users aged 18 and above).

Table 3. Partial correlation ($n = 15$)

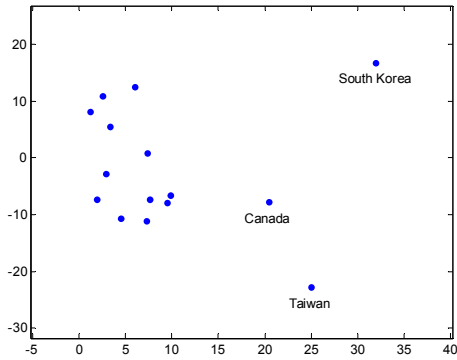
CD sales - Downloads 18+



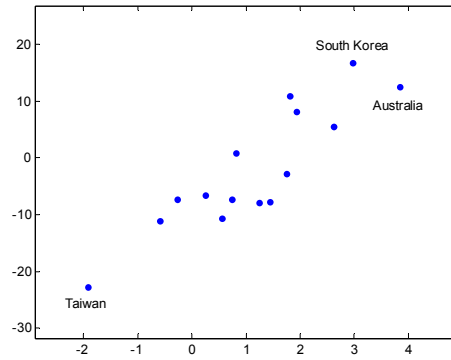
CD sales - Downloads 18-25



CD sales - Broadband penetration



CD sales - growth



2.2 Model

Let $y_{i,t}$ denote the sales of CDs in units (in log) in country i at time t . $w_{i,t}$ is the variable that captures internet piracy. $x_{i,t}$ is a vector of control variables in country i , at time t and β is the set of associated coefficient. The unobservable variable is denote $\varepsilon_{i,t}$ assume to have a zero mean and independent of the vector of explanatory variables. To eliminate country specific effects, we write the model in first difference:

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

where α and β are parameters to be estimated and $i = 1, \dots, n$, $t = 2001$. In particular, α is the elasticity of CD sales to online piracy.

However we do not have online piracy variables in difference but rather in level. Thus we estimate the model:

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

where w_i^* are our 2002 data on music downloads (and 2001 data on broadband penetration).

The implicit assumption is that the elasticity calculated using the cross-section dimension is the same as the elasticity obtained using first-differences: $\alpha = \alpha^*$. In order to obtain the net effect of online piracy on Cd sales, we have the reconstruct the increase in online piracy in each country.

2.3 Results

Table 4 gives our estimation results. We have run separate regressions including growth and each specification of online piracy. We make some preliminary remarks. We also included separately a constant, MC sales, and CD penetration. This resulted in insignificant coefficients and no increase in R^2 . We also interacted downloads with broadband penetration as such and then divided by internet penetration, with unsatisfactory results. We run regressions for the period 1999-2001 with similar qualitative results. The fit was however poorer as the period extends away from our recent data on broadband and music downloads.

Table 4. Estimation results 2000-2001,

	(1)		(2)		(3)		(4)		(5)	
	coef.	s.e.	coef.	s.e.	coef.	s.e.	coef.	s.e.	coef.	s.e.
Growth	4.42	1.14	6.44	0.91	6.50	0.88	5.98	1.02	6.21	0.97
Broadband	-0.42	0.16								
Downloads 18-25			-0.15	0.03						
Downloads 18+					-0.24	0.04				
Internet*Downloads 18-25							-0.27	0.06		
Internet*Downloads 18+									-0.46	0.09
R^2		0.52		0.78		0.79		0.71		0.74

We now comments on our estimation results. First, contrary to Liebowitz (2002) we find a strong positive effect of income (GDP) on CD purchases. Secondly, broadband

adoption times differ across countries; this could explain the poor fit in specification (1). The best linear regression comes from specification (3) where the increase in music downloading is proportional to the percentage of internet users who have ever downloaded. Indeed, the true measure of internet piracy is the number of heavy music downloaders. Comparing to specification (5), we know from the analysis of the previous section that the number of daily music downloaders is proportional to the share of internet users who download music and is less related to internet penetration. Moreover, Downloads number concern a subpopulation of adults who are more than 18, while internet penetration is defined in percentage of total population.

Based on the estimated coefficients, we can assess the importance of internet piracy on the decline in CD sales. In specification (3) the implied effect of music downloading is $.24 \times .075$ (music downloads increased by 33% each year in the U.S. and is assumed to have done so in other countries) = .018 which represents 28% of the 6.65% drop in CD sales in the US. In the next section, we will present factual data on music downloading in the U.S. in order to make sense of this percentage. Using specification (1), this number decreases to $.42 \times .033 = .014$ (broadband also roughly increased by 33% during the period) or a little bit more than 21% of the decrease in CD sales.

Averaging across countries, we find that internet piracy can be associated with an (unweighted) average 1.76% drop in CD sales worldwide (last entry of column 1 in Table 5). We also compute the loss in percentage of the CD sales decline across countries (third column) and find that on average internet piracy explain 25% of the CD sales (in absolute value) in the top 10 markets for recorded music.

Table 5. Implied losses (2000-2001)

Country	Implied losses	CD Sales	Implied losses /CD sales
USA	-1.84	-6.65	0.28
Japan	-0.92	-11.20	0.08
UK	-1.29	8.10	-0.16
Germany	-1.56	-10.73	0.15
France	-1.56	10.78	-0.14
Canada	-2.02	-7.76	0.26

Mexico	-1.61	-7.32	0.22
Spain	-1.79	5.49	-0.33
Italy	-2.11	-2.85	0.74
Australia	-1.70	12.45	-0.14
Netherlands	-1.84	-7.93	0.23
Sweden	-1.84	0.76	-2.41
South Korea	-2.30	16.66	-0.14
Belgium	-1.70	-7.44	0.23
Taiwan	-2.34	-22.87	0.10
Average	-1.76		

3 Interpretation of implied loss in the U.S.

3.1 Summary of facts on music downloading

Survey data gathered for the period 2000-2002 suggest the following stylized facts for the U.S. that are documented in the Appendix

1. 138 million of adults aged 12 and older had access to the internet in 2000
2. about 90% of internet users have access at home, while 10% can only access from work, colleges or other places in 2001-2002
3. There are two types of music downloaders: heavy internet users have a fast connection and 15% of those download on any given day; only 5% of light internet users download on any given day (in 2001). Similarly, about 2/3 of music downloaders have stored a relatively small number of files (on average 13 in 2000); 1/3 of music downloaders store an extremely large number of files (on average 258 in 2000).

3.2 Assessing “internet piracy”

Our analysis of the downloading phenomena suggests that there are two types of music downloaders: heavy and light. Only heavy music downloaders can potentially negatively influence CD sales, as there is a great likelihood that light users download music for sampling purpose only. Here we make a list of assumptions:

1. Internet users who download music purchase on average the same number of CD as the aggregate population = 3.7 per capita in 2000, multiplied by a factor (1,

- 1.15, 1.3, 2, 4) that takes into account above average CD consumption of music downloaders, burning and sharing with friends and family members.
2. Only users who download on a regular (daily) basis are a threat to copyright owners.
 3. These same users reduce their net CD demand by 10 to 100%.
 4. Work internet users have the same downloading pattern as broadband internet users and college students.
 5. Home internet users have the same music downloading behavior regardless of whether they only access from home or also from work.

Assumption 1 can be defended on the ground that entertainment time is inelastic. Indeed, demand for music does not significantly increase or decrease with income and age.

Assumption 2 is reasonable given the numbers presented in the tables above. Assumption 3 (a) is heroic as there is hard to believe that there is full substitution between mp3 and Cds for heavy music downloaders who should also have a higher willingness to sample and to pay for the original. Assumptions 3 (b)-(c) describe weaker substitution patterns between CDs and Mp3. Assumption 2, 5 and 6 are made by lack of precise information on these types of internet users

We know that 90% of internet users can connect online from home. Among these users, 25% have broadband and 75% do not. Moreover roughly 15% broadband users download music on any given day, while 5% of dialup home internet users do so. Internet users who only download from work or other places (including college students) represent 10% of the online population and are assumed to have the same behavior as broadband users: 15% download music on a regular basis. To sum up, among home users, 7.5% ($=.25 \times .15 + .75 \times .05$) download music on a regular basis. On the other hand, 15% of work users do so. Thus the total percentage of internet users who are potential copyright “offenders” is 8.25% ($= .9 \times .075 + .1 \times .15$) of the online adult population aged 12 and above ($= 138$ millions) = 11.385 million. Suppose that they reduce their past consumption by 20% (as stated in assumption 3 (c)) and that there are no multiplier effects due to sharing and above average CD consumption of downloaders (second column). Then the lost sales due to online downloads is 8.43 million of units ($= .0825 \times 3.7 \times .20 \times 138$) or

14% of the 60 millions drop in CD sales between 2000 and 2001. Suppose full substitution, no sampling and no multipliers: lost sales increase to 42.12 million or about 70% of the drop of total CD sales. Table 6 summarizes simulation of lost sales due to piracy for different types of substitution and Mp3 and different multipliers effects. The percentage lost of .28 that was implied by our estimation results is compatible with a substitution rate between Mp3 downloads and CDs that ranges between 10% to 40%, for different values of the multiplier effect.

Table 6. Possible effects of Mp3 on CD sales for 2000-2001 (% of lost sales due to music downloading, 2000-2001)

Substitution between MP3 and CDs	Multiplier effect	1	1.15	1.35	1.6	2	4
1	0.70	0.81	0.95	1.12	1.40	2.81	
0.9	0.63	0.73	0.85	1.01	1.26	2.53	
0.8	0.56	0.65	0.76	0.90	1.12	2.25	
0.7	0.49	0.57	0.66	0.79	0.98	1.97	
0.6	0.42	0.48	0.57	0.67	0.84	1.68	
0.5	0.35	0.40	0.47	0.56	0.70	1.40	
0.4	0.28	0.32	0.38	0.45	0.56	1.12	
0.35	0.25	0.28	0.33	0.39	0.49	0.98	
0.3	0.21	0.24	0.28	0.34	0.42	0.84	
0.25	0.18	0.20	0.24	0.28	0.35	0.70	
0.2	0.14	0.16	0.19	0.22	0.28	0.56	
0.1	0.07	0.08	0.09	0.11	0.14	0.28	

4 Conclusion

The main source of the decline in CD sales between 2000 and 2001 is due to a negative economic environment. Using cross-section regressions, we estimate that Internet piracy can be associated with a decline of 1.76% in CD sales worldwide between 2000 and 2001. There are reasons to believe that the effect of internet piracy is likely to decrease over the coming years. Indeed, there is self-selection in the adoption of broadband and P2P technologies, so that users who have the highest propensity to download music adopt first. Given our factual data on music downloads in the U.S.,

record companies missed opportunities to generate revenues from broadband and P2P users who have the highest propensity to sample and purchase online, either through ads (internet streaming) or through direct purchase (lack of business model for mp3 downloads).

5 References

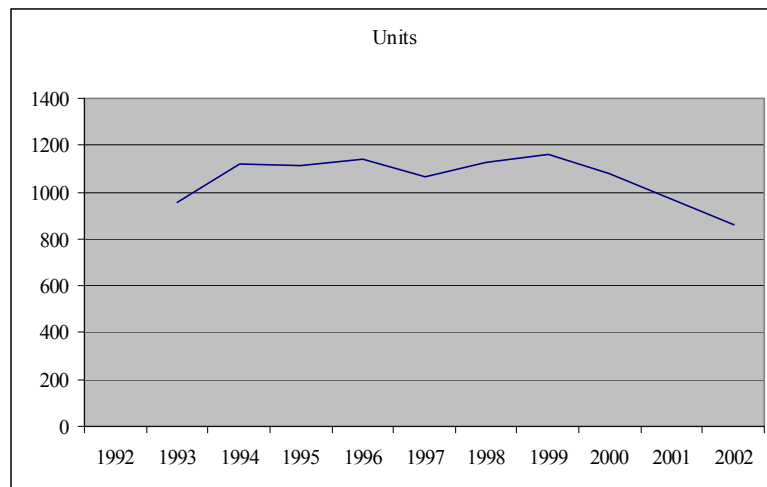
- Duchene, A. and P. Waelbroeck (2003), "Does P2P Harm Copyright Owners? Distributing and Protecting Digital products", CERAS, mimeo
- Hui, K. L., and Png, I. P. L., 2002, "Piracy and the Legitimate Demand for Recorded Music", Mimeo.
- Liebowitz, S., 2002, "Record Sales, MP3 Downloads and the Annihilation Hypothesis," Mimeo, U.Texas, Dallas.
- Peitz, M. and P. Waelbroeck (2003a), "Piracy of Digital Products: A Critical Review of the Economics Literature," Mimeo
- Peitz, M. and P. Waelbroeck (2003b), "An Economist's Guide to the Technology, Law and Business of Music Distribution," Mimeo
- Peitz, M. and P. Waelbroeck (2003c), "Making Use of Online File-sharing in Music Distribution," Mimeo

6 Appendix: empirical observations for the U.S.

6.1 The market

We start by looking at detailed data available for the United States between 1993 and 2002. The market was relatively flat between 1993 and 1999 and there is currently a sharp decline starting from 2000 (Figure 1).

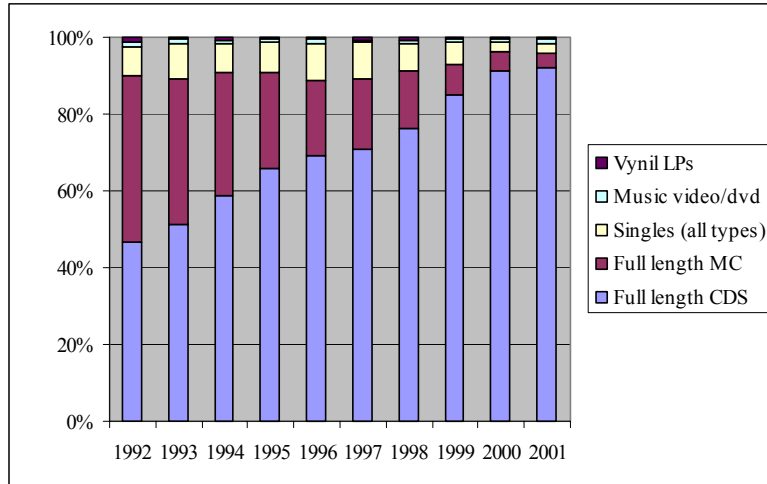
Figure 1. Total retail Units (CDs, Single, Cassettes, LP, Music Video)



Source: RIAA (in millions)

However, several factors in addition to the global negative economic environment can have decreased music sales around that period: substitution between music media and substitution with other forms of entertainment. Figure 2 indicates that the year 2000 also coincides with the end of a strong substitution/replacement effect between cassettes and CDs.

Figure 2. Substitution between different types of media

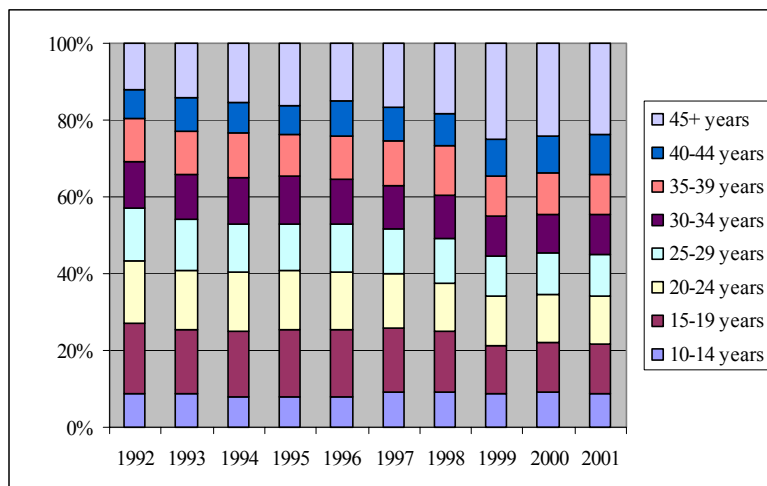


Source: RIAA and own calculations

In addition, it turns out that different cohorts of consumers do not have the same purchasing habits over time. This reflects the fact that the older population is replacing its musical cassette/vinyl collection with CDs, while the younger population has the possibility to consume new types of entertainment products such as video console games, internet chat and mobile phones (all new during the period). These two trends are apparent from Table 7, much prior to the CD sales decline of 2000, although they are flattening over the last couple of years.³

Table 7. Consumer profile

³ The telephone survey is based of past month buyers (over 3000 per year). The data is weighted by age and sex and then projected to reflect the U.S. population age 10-and-over.



Source: Peter Hart Research for the RIAA and own calculations

We make two preliminary remarks before looking in more details at the losses due to online piracy and the sampling effect of digital copies. First, we believe that while the major record companies have launched official online distribution sites, no good business models have been introduced so far.⁴ Thus, the sampling effect can only occur through third parties, that is, by downloading from unauthorized file-sharing technologies. If sampling occurs, consumers purchase CDs of music that they like in record or other stores. This means that they do not fully substitute CDs by MP3 files.

Secondly, to have a positive or a negative effect, P2P technologies must significantly affect the population of potential buyers. In particular, in order to claim losses due to internet piracy, one needs to argue that the users of P2P technologies are precisely those who would have purchased the original at its full retail price. Before providing our point of view on this issue, we first need to understand who copies what and how much copying actually occurs.

6.2 Who copies?

We first assess the number of online users and how they access the internet.

⁴ See "Face the (Digital) Music," Wall Street Journal, Dec. 2, 2002. See also our discussion in Peitz and Waelbroeck (2003b).

Table 8 shows the evolution of the total online adult population aged 18 and older, which was about 63% in 2000.

Table 8. US adults online

Year	Adults online (%)	in millions
2002	66	137
2001	64	127
2000	63	121
1999	56	113
1998	35	70
1997	30	59

Source: Harris Interactive, February-March 2002

The relevant number to analyze music download is however the percentage of adults aged 12 and older. Using the 2000 US census, IPSOS-REID evaluates this population to 218.6 millions. If we assume that the same percentage of 12-17 teenagers are online as the other adult population, the total number of adults aged 12 and over who are online equals 138 (=218.6*.63) million in 2000. We will use this number in our preliminary attempt to assess the role of internet piracy on CD sales at the end of this section.

Next, we look at how consumers get access to the internet. Table 9 decomposes the spectrum of internet access by the type of home users: heavy users have broadband connection, while light users alternate between being connected or not; similarly, net dropouts stopped to use the internet for some reasons, while net evaders stay disconnected on purpose. This table shows that average and heavy (broadband) users account for 30-42% of the population.

Table 9. Spectrum of internet access (in percentage of total population)

Heavy	Average	Light	Total online	Net dropouts	Net evaders	Truly unconnected	Total offline
13	17-29	16-28	58	10	8	24	42

Source: PEW Internet Report, Mar.-May and Dec. 2002 surveys

Table 10 shows that music downloading is more popular among users aged 12-34 (especially among 12-17), although they do not represent the largest percentage of the population.

Table 10. Profile of music downloaders

Have you ever downloaded a music or MP3 file off of the internet from your PC? (in percentage of category)	
Total (1112 obs.)	23%
Gender	
Male (566 obs.)	26%
Female (546 obs.)	20%
Age	
12-17 (111 obs.)	44%
18-24 (125 obs.)	42%
25-34 (194 obs.)	35%
35-54 (394 obs.)	18%
55+ (282 obs.)	5%

Source: Ipsos-insight, Tempo: Keeping pace with online music distribution, 2001

To have a better understanding of who downloads music, it is useful to analyze how music downloaders access the internet. According to Pew Internet Report, Getting serious online, 49% of internet users have only access at home, 8% only access at work, 39% have access at both home and work and 4% have access in other places such as school and library. So, roughly 90% internet users have access at home, while 10% have only access from work, college or other places.

6.2.1 Home internet users

We can classify home internet users by the type of connection they have. Table 11 describes profiles of broadband users and dial-up users. Heavy internet users (Broadband) are highly educated, wealthy people aged 25-45 living in a suburban area.

Table 11. Internet user profile by type of internet connection (% of respective population)

	Home	Dial-up
--	------	---------

	broadband (n = 507)	(n=1391)
Sex		
Male	56	49
Female	44	51
Community type		
Rural	11	24
Suburban	60	51
Urban	29	25
Age		
18-24	11	15
25-34	20	22
35-44	28	27
45-54	26	21
55+	14	15
Education		
High-school or less	13	37
Some college	28	28
College grade or more	59	35
Race		
White (not hispanic)	85	79
Black (not hispanic)	4	8
Hispanic	5	9
Other	6	5
Income		
Less than \$30K	6	17
Between \$30K and \$50K	14	24
Between \$50K and \$75K	20	19
More than \$75K	45	24

Source: PEW internet tracking survey February 2002 (broadband) and August-September 2001 (dial-up)

Now who downloads music? Table 12 shows that, while the whole population has embraced the internet revolution, the subpopulation of downloaders represents less than 15% (58% internet users* 25.5% ($=\frac{22+29}{2}$ on average)) of the population. However, this number increased to 20% in August 2002. Now music downloading is in proportion highest among poor non-white and young internet users with little education. However, they represent an extremely small proportion of heavy internet users as documented in Table 11. Moreover, it is hard to argue that these users are causing any harm as they are the less likely to purchase the original at the full retail price.

Table 12. Downloading profiles: internet users who ever downloaded music over the internet (in percentage of the population of internet users)

	July-August 2000	February 2001
All Adults	22	29
Men	24	36
Women	20	23
Whites	21	26
Blacks	29	30
Hispanics	35	46
<hr/>		
Age cohorts		
18-29 years	37	51
30-49 years	19	23
50+ years	9	15
<hr/>		
Household income		
Under \$30000	28	36
\$30000-\$50000	24	31
\$50000-\$75000	20	29
\$75000+	15	24
<hr/>		
Education		
Less than high school	38	55
High school	25	31
Some college	25	32
College degree or more	15	24
<hr/>		
Internet user experience		
Less than 6 months	20	27
6 months to 1 year	20	25
2 to 3 years	24	28
3 or more years	22	33

Source: PEW Internet Tracking Report, April 2001

In table 12, we note that for income and internet experience, the slopes are negative as expected but not very significant. Table 13 provides more recent data on music downloading by age.

Table 13. Percent of US home internet users who have ever downloaded and stored mp3 files by age

18-24	81
25-34	62
35-44	50
45-54	40
55-64	24
65+	13

Source: Parks and Associates, April 2002

Now, these two tables do not control for how much downloading is going on. It can well be that a large percentage of the population of internet users very occasionally downloads music, while a very small number of internet users make most of the downloading. We will come back to this issue later on.

6.2.2 College students

Harris Interactive/360 Youth Fall 2002 Study finds that 93 percent of college students access the internet in a given month, 88% own a computer and 56% have a broadband connection. They spend more than \$210 billion in sales last year. Moreover, 2/3 of college students have paying jobs that represent \$53.9 billion in discretionary spending annually. Most of the spending goes on entertainment and leisure related expenses. College students were spending \$5 billion on travel, \$790 million at the movies, \$390 million on attending music concerts, \$318 million at amusement parks and \$272 million at professional sporting events.

This pattern is confirmed by Pew Internet Project (1021 college students March-June 2002). They were 14.5 million students enrolled in colleges and universities in 2002 or 5% of the population. College internet users who have ever downloaded music is larger than the average internet population (60% have done so compared to 28% overall) and three times as likely to download music on any given day (14% compared to 4% overall, a percentage similar to the respective percentage for broadband users). College students also lead other internet users in file sharing of all kind (44% against 26% overall). Moreover they share files other than music in a greater proportion: 52% downloaded files other than music compared to 41% for the overall internet population.

Finally, according to comScore Networks, 7.7% of US internet users connect from college and university based PCs in 2001.

6.2.3 Work users

According to Pew Internet Report "Getting serious online," in January 2002, 55 million people went online from work compared to 43 million in March 2000. Some 55% of internet users with access at work went online on any given day in 2001, up from 50% in 2000 (more internet access from work).

6.3 What and how?

Table 14 reports the most common activities of internet users: looking for information on products, jobs, travel, hobbies, travel and entertainment. To get a preliminary order of magnitude, music listening/downloading is as popular as checking sports scores and playing games online.

Table 14. Internet activities: how they changed over the course of a year (in percentage of internet population)

Millions who have ever done	March 2001	March 2000
Communication activities		
Email	100	75
Instant messages	48	36
Fun activities		
Info on a hobby	83	64
Online just for fun	66	53
Video/audio clips	56	40
Listen/download music	40	30
Play a game	40	30
Sports scores	38	28
Information utility activities		
Product information	82	63
Travel information	72	55
Information on movies, books, music	69	53
News	64	52
Health information	64	47
Government website	60	42
Job-related research	52	41
Financial information	45	38

Look for job	44	31
Look for place to live	32	20
Religious/spiritual information	27	18
Transactions		
Buy a product	58	40
Buy a travel service	46	29
Online banking	25	14
Online auction	22	12
Buy/sell stocks	13	10
	n=862	n=723
	internet	internet
	users	users

Source: Pew Internet Report, Getting serious online, March 2002

Table 15 highlights the prevalence of audio streaming and internet radio among online music activities.

Table 15. Music related internet activities

Have you ever (in percentage of all respondents, 1112 obs.)	
Downloaded music or mp3 files off the internet	23%
Listened to internet radio	27%
Listened to streamed music clips or audio files	21%
Listened to prerecorded CD in you PC Cd-Rom drive	37%

Source: Ipsos-insight, Tempo: Keeping pace with online music distribution, 2001

File-sharing is popular among internet users aged 24 or less. But again they represent a minority of the population. It is not clear that the current number of P2P users will continue to dramatically increase as illustrated by the flat trend in Table 16. It is also worth noting that the large proportion of P2P users have less than 17 and do not represent a significant percentage of the population with large purchasing power.

Table 16. File-sharing

Have you downloaded digital music files (or MP3 files) from an online file-sharing service (such as Moprheus, Napster or Kazaa)?	April 2002		December 2002	
	Ever	in the past 30 days	Ever	in the past 30 days
Total (1112 obs.)			19%	9%
Gender				
Male (566 obs.)			26%	13%

Female (546 obs.)			12%	6%
Age				
12-17 (111 obs.)	42%	23%	52%	32%
18-24 (138 obs.)	45%	26%	44%	24%
25-34 (181 obs.)	26%	11%	23%	8%
35-54 (394 obs.)	14%	4%	12%	5%
55+ (282 obs.)	2%	1%	3%	1%

Source: Ipsos-insight, Tempo: Keeping pace with online music distribution, 2002

To assess how much sampling and freeloaders are out there, we provide in Table 17 these respective numbers.

Table 17. Music sampling 1/4

Did you buy the music you downloaded or did you get it for free?	%
Bought it	15
Got it for free	79
Don't know/refused	9
<hr/>	
Did you download music that you already own on a CD or tape or did you download new music?	
Music already owned	28
New Music	63
Don't Know/refused	9
<hr/>	
After you downloaded music to your computers and listened to, how many times if ever have you bought that same music on a CD or cassette	
Most of the time	21
Some of the time	29
Only a few time	19
Never	26
Don't know/refused	5

Source: PEW internet tracking, July-August 2000.

There are only 26% of freeloaders in the survey, while 50% of music downloaders have declared to have actually purchased the original on a regular basis. However, these numbers could be biased as they are self-reported in a strengthening legal environment.

Table 18. Sampling 2/4

What type of music have you downloaded?

	Yes	No	Don't know refused
Music you'd heard before by artists you were familiar with	86	9	5
New music by artists you were already familiar with	69	27	4
Music by artists you had never heard before	31	65	4

Source: PEW internet tracking, July-August 2000.

Table 18 hints at the fact that 31% of music downloaders have listened to new artists (this is a lower bound on sampling effect). This is somewhat confirmed by numbers reported in Table 19 and 20.⁵

Table 19. Sampling 3/4

"has the genre of music that you typically listen to/or purchase changed since you initially began downloading music or mp3 files off o the internet? (834 obs.)

No	71%
Yes	30%

"In what ways?" (242 obs.)

was able to experiment with different forms of music	27%
like different/a range of music	23%
introduced to new age/techno/electronica	10%
more aware of new bands, groups, artists, songs	10%
listening to more country/introduced to country	6%
listening to more classical/introduced to classical	5%
listening to more pop/introduced to pop	4%
listening to more hip-hop, rap/introduced to	4%
listening to more jazz/introduced to jazz	4%
other	19%

Source: Ipsos-insight, Tempo: Keeping pace with online music distribution, 2002

Table 20. Sampling 4/4

⁵ Recent theoretical contributions highlight the positive role that sampling can play for profits and welfare. See Duchene and Waelbroeck (2003) and Peitz and Waelbroeck (2003c).

Thinking of the music that you either downloaded or streamed off of the internet, or the internet radio stations you have listened to off of the Internet, would you say you have done this for free, on a fee-basis either by subscription or per-use, or a mixture of both (834 downloaders)

Only on a fee-paid basis	1%
Only for free	92%
Both	7%

"And how likely would you be actually pay to download or stream music off of the internet if there was no free material available" (769 downloaded for free)

likely	1%
somewhat likely	11%
not likely	84%

"Since you initially began downloading music or mp3 files off of the internet, would you say that your Compact Disc purchases have" (834 obs.)

Decreased	19%
Increased	24%
Stayed the same	57%

Source: Ipsos-insight, Tempo: Keeping pace with online music distribution, 2002

6.4 How much?

The fact that the majority of internet users who ever downloaded music only occasionally do so is illustrated by Table 21. Here we decompose downloading activity by internet connection on a daily basis.

Table 21. Daily internet activity by internet connection (in % of respective category)

	n=507	Broadband elite (25%)	Other broadband (75%)	n=1391
	All Home broadband			Dial-up
Communications				
Email	67	58	80	52
Instant messaging	21	48	14	14
Chat rooms	10	23	7	5

<hr/>				
Information seeking				
News	46	56	49	24
Job-related search	36	64	32	14
Look for product information	32	68	24	18
Research for school or training	24	50	19	9
Look for travel information	23	61	14	6
Look for medical information	21	52	13	8
<hr/>				
Information producing				
Share computer files with others	17	50	8	4
Create content (e.g. web pages)	16	38	10	3
Display/develop photo	14	45	5	1
Store files on the internet	8	21	4	n/a
<hr/>				
Downloading				
Download games, video, pictures	22	61	12	4
Download music	17	43	10	6
Download movie	5	17	2	n/a
<hr/>				
Media/streaming				
Watch video clip	21	55	12	6
Listen to music/radio station	19	48	11	4
Watch movie	4	12	2	n/a
<hr/>				
Transactions				
Online banking/bill paying	22	39	19	6
Buy a product	21	59	11	3
Buy a travel service	14	47	5	2
Auction	10	28	5	3
Buy groceries	5	18	2	1
Buy/sell stocks	5	17	2	1
Gamble	2	7	1	n/a
<hr/>				
Entertainment activities				
Hobby information	41	69	37	18
Browse just for fun	39	55	39	21
Play a game	22	46	17	10
Visit adult web site	6	17	3	1

Source: PEW internet tracking survey February 2002 (broadband) and August-September 2001 (dial-up); elite broadband users are doing on average 10 or more activities on a daily basis and represent 25% of the broadband population.

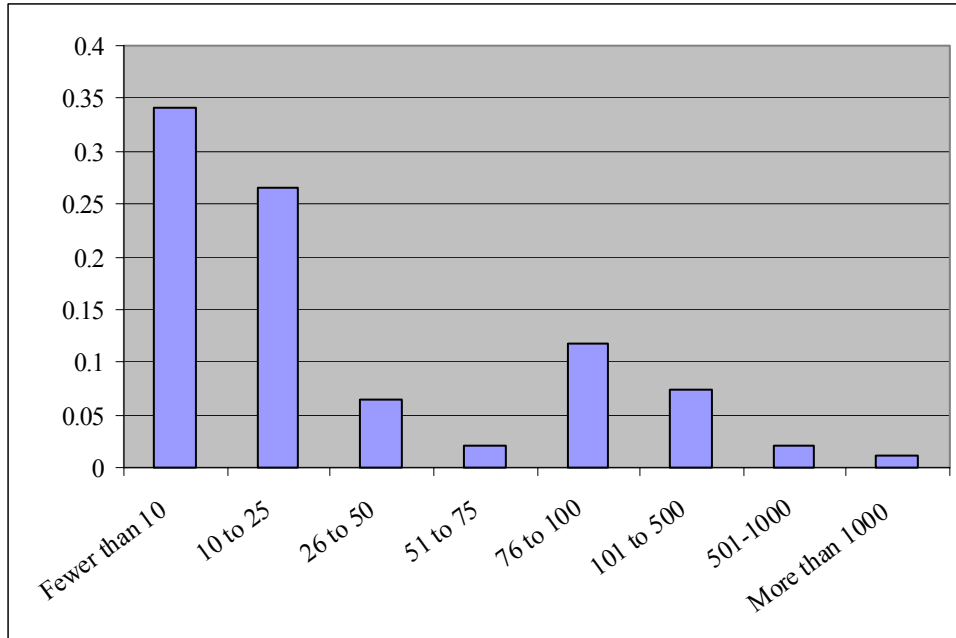
Now, looking at the home users who download on a daily basis (i.e. those who could really potentially influence CD sales), it turns out that their number as a percentage of the population, 4.9% (2.7% (= .06*.45 of dialup users) + 2.2% (= .17*.13 of the broadband users), is relatively low. Moreover, those who actually share and download music are in fact elite broadband users and represent 1.4% (= .43*.25*.13) of the population. But, as illustrated in Table 11, these users are also the most likely to purchase the original as they

have no credit constraint and are not afraid to use the internet to make purchase. Thus the net effect on CD sales is ambiguous. Moreover, it is also worth noting that heavy internet users listen to streaming audio as much as they download (19% vs. 17%). In the former case, copyright owner collect royalties and it is difficult to assess the net contribution of broadband users to revenues of big label companies, especially because it is difficult to obtain information on how much money is made from streaming audio.

Figure 4 illustrates the distribution of users who have downloaded any given number of files as a percentage of the total number of music downloaders (238 out of the 2109 who answered the survey or 11.2% in 2000). This distribution seems to be bimodal, as many as 62% of users who downloaded music have less than 25 music files on their computers (about 1.6 of a CD, if a CD includes 15 songs). However 23 % of music downloaders have more than 75 files on their computers, while few consumers download an average. It is hard to argue that light downloaders are causing any harm to big label companies. The only serious threat comes from average and heavy downloaders (with 50 files or more) who represent 4.6% ($=.15*.31$) of the population on average during 2000-2001, a percentage similar to the one found using daily downloads. Note that all these numbers are meaningful as they were collected a year after the creation of Napster.

Taking the average across all categories yields 313 songs per computer. Figure 3 also suggests that there are two types of music downloaders with two different types of cumulative distribution: one asymmetric with a mode at 0 and one unimodal with a mode around 100 and a mean around 500.

Figure 4. Number of files downloaded (in percentage of the population who ever downloaded music)



Source: PEW internet tracking, July-August 2000.

The following table provides a breakdown of music files stored on computers by age.

Table 22. Average number of mp3 files on home computers by age

65+	72
55-64	124
45-54	177
35-44	340
25-34	721
18-24	348

Source: Parks and Associates, April 2002

The average number of files is 297 per computer. Although these numbers could sound alarming for copyright owners, we know from Figure 4 that only a few number of internet users download a lot.