

Measuring the Effect of Online Music Piracy on Music Sales

Alejandro Zentner*

University of Chicago

Preliminary Version

Abstract

Online piracy may substantially undermine intellectual property rights of digital goods. This paper concentrates on the music industry. I estimate the effect of online music piracy on music sales using two databases: a European individual level cross section of more than 15,000 people from October 2001 and a panel of aggregate music sales by country for years 1997-2002. In the micro-data, a simple comparison of means shows that people who regularly download music online are more likely to buy music. The positive relation persists when controlling for many individual level characteristics. However, simultaneity between tastes for music and peer-to-peer usage makes it difficult to isolate the causal effect of online piracy on music purchases. To break that simultaneity, this paper uses broadband connection and measures of internet sophistication as instruments for peer-to-peer systems usage. The results suggest that, for the group of users of peer-to-peer systems, piracy reduces the probability of buying music by 35% to 65%. Based on my estimates, back of the envelope calculations indicate that online music piracy may explain a drop in music sales of 7.8% to 14.5%. Using the number of internet users by country as a measure of users of P2P systems, the panel of aggregate data shows a larger impact of piracy on music sales.

* Department of Economics, University of Chicago (alezentn@uchicago.edu). I wish to thank Austan Goolsbee for his advice and encouragement. I also thank very useful comments from Gary Becker, Dennis Carlton, Carolina Czakiewicz, Julio Elias, Thomas Hubbard, Anupam Jena, David Levine, Steven Levitt, Jose Liberti, Douglas Lichtman, Laura Martinolich, Kevin Murphy and Chad Syverson.

I- Introduction

The global music industry was quite successful during the 1990s. According to the International Federation of the Phonographic Industry (IFPI), album sales grew from US\$24.1 billion to US\$38.6 billion during the decade. Those days are over and the industry is now struggling. Sales have been falling for the last three years: global sales dropped 7.1% (to US\$30.9 billion) in 2002, 8.8% in 2001, and another 5.0% in 2000.

This downturn coincides with the proliferation of online music file sharing. In June 1999, Napster was created and made the work of many artists available for free. Its popularity was immediate. According to Mediаметrix, a company that provides internet rankings and measurement, Napster was the fastest software adoption in history (CNN, 2000). Given the impact, the Recording Music of America (RIAA) soon filed a motion against Napster in the U.S. District Court of San Francisco for “engaging in or enabling, facilitating or assisting others in the copying, downloading, uploading, transmission, or distribution of copyrighted musical work or sound recordings protected by copyright or state law without the express permission of the rights owners” (US District Court, 2000). Napster was shut down in February 2001. However, many peer-to-peer (P2P) alternatives to swap music over the internet remain available. KaZaA, a Napster successor, holds the new record of most downloaded software with more than 230 million users worldwide (KaZaA, 2003). In all the countries considered by the Yahoo Buzz Index, an index that measure internet search using the Yahoo search engine, KaZaA was the number one search term on the internet in many weeks during 2003.

File sharing has an important online presence. In June 2001, IFPI estimated 3 million simultaneous global users and 500 million files available for copying at any given time. In 2002, the IFPI estimate went up to 5 million simultaneous users and 900 million files (IFPI, January 2003). In 2002, NetPD, a company that provides protection services to copyright owners whose material is being pirated through the internet, reported that 3.6 billion files were downloaded “monthly,” of which around 60% and 70% were music files (The Observer, 2002). Most popular albums are available for online swapping almost immediately after release, and in some cases, such as Oasis and Eminem, even before (IFPI, 2002). Copy protection technology has been ineffective.

The development of broadband facilitates music swapping. A soundtrack that takes more than 12 minutes with a dial-up connection can be downloaded in as fast as 20 seconds with a high-speed connection. Universities have very fast connections and Napster and its successors were banned in many of them because file swapping was consuming much of the available bandwidth. In the case of the University of Illinois at Urbana-Champaign, this amounted to 75% of the total bandwidth (BusinessWeek, 2000).

Online piracy is not exclusive to music. Other digital copyrighted goods, such as movies, software, games, books, etc, are also being pirated. The development of fast connections is extending piracy to digital goods with high content. Some top movies are available to swap online on theater's opening week.

What is the magnitude of the impact of file sharing on music sales? Is file sharing responsible for the drop in sales during the last years? Inside the music industry, it is generally accepted that music sales have been affected by file sharing. But there has not been much empirical work done to measure the extent of this effect.

The question is important because file-sharing technology "has substantially undermined the effective protection offered by copyright" (Romer, 2002). Strong property rights create monopoly distortions but weak property rights may lead to low creation of artistic work or innovation.¹ The balance between these opposing forces has not been much informed empirically. Knowledge of the actual estimates of the impact of file sharing is essential in the determination of the strength of property rights of digital goods.

To the best of my knowledge, the only empirical attempt to measure the effect of online file swapping on music sales is the one prepared in the legal battle against Napster.² To measure this effect, the plaintiff hired Soundscan, a company that developed an information system to capture point-of-sale data on music sales in more than 18,000 stores throughout the US. Soundscan compared sales means for the first quarter of years 1997, 1998, 1999 (when Napster was not available) and 2000, of all stores, stores within one mile of any college or university, stores within one mile of the top 40 most wired

1 See Boldrin and Levine (2002) and Klein, Lerner and Murphy (2002) for a discussion of the effect of file sharing on copyright value.

2 Kai-Lung and Png (2003) analyze the effect of comertial piracy on music sales with data from the years 1994-1998.

universities and stores near universities that banned Napster use after the first quarter of 2000. From the first quarter of 1999 to the first quarter of 2000, national sales grew 6.6%, sales near all universities dropped 2.6%, sales near most wired schools dropped 6.2% and sales near schools where Napster was banned after the first quarter of 2000 fell 8.1%. However, sales near universities were falling since 1998, at a time when Napster was not available and in which national sales were growing, casting doubts on the conclusion of Soundscan's report (Fine, 2000).

The objective of this paper is to obtain estimates of the causal effect of online file sharing on music sales. In the microdata, the main difficulty in the estimation of this effect is the presence of unobserved heterogeneity in music tastes among individuals. To control for this heterogeneity, it would be ideal to follow individuals through time. In the absence of such a panel data, I use a major European cross section database from October 2001. With this data, I attempt to identify the causal effect of downloading MP3 files on the probability of buying music.

Ordinary least squares shows a positive -- though not significant -- relationship between peer-to-peer usage and the probability of purchasing music. However, OLS is biased and inconsistent in the presence of unobserved heterogeneity. People who download music have a higher taste for music. To overcome this obstacle, my approach is to use instrumental variables techniques. It is required an instrumental variable that is correlated with peer-to-peer systems usage, but is otherwise unrelated to music purchases. In this paper, I employ broadband connection and measures of internet sophistication as instruments. My estimates indicate that peer-to-peer usage reduces the probability of buying music between 35% and 65%.

Given the dramatic change in the estimated effect in the instrumented regressions, special attention is given to the validity of the instruments and robustness of the results. First step regressions show a positive and significant correlation of the instruments with peer-to-peer usage. Test of overidentifying restrictions are consistent with the exclusion of the proposed instruments from the music purchase equation. I test the robustness of the results by checking the possible presence of regional unobservables, by restricting the analysis to internet users and to individuals who had internet connection before Napster started, and by analyzing how MP3 downloading affects purchases of other goods. I also

analyze the relationship of the instruments with variables that may signal taste for music and value of time.

With the level of people who downloaded MP3s in 2002, and assuming that the purchased *quantity* conditional on purchasing is the same for users and non-users, my estimates imply a reduction of music sales units in the countries considered of between 7.8% and 14.5%.

The individual level database does not contain quantities of music purchased; only allowing back of the envelope calculations to predict the effect of online piracy on music sales. To analyze the effect on quantities, and as an additional check of the effect of online piracy, I combine a panel of aggregate music sales by country for the years 1997-2002 with data on internet usage by country as a measure of the number of P2P users. Turning to the aggregate data implies working with a fewer number of observations and controls but it allows checking the effect on quantities. Controlling for aggregate level variables, I find that the country level of peer-to-peer users significantly affects the change in music sales pre and post existence of P2P systems. The estimated effect with the aggregate data shows an even larger effect of online piracy on music sales.

The paper is organized as follows. Section II summarizes the music industry. Section III describes the micro-data. Section IV presents the empirical strategy and results using the micro-data. Section V discusses the implications of the estimated coefficients over music sales. Section VI presents the empirical analysis with the aggregate data. Finally, section VII concludes.

II- The Music Industry

Music sales in 2002 were US\$32.2 billion.³ 41% of these sales were made in North America, 34.5% in Europe, 18.6% in Asia -- with Japan representing more than 80% of Asian sales --, 3.1% in Latin America, and 2.7% distributed among Australasia, the Middle East and Africa. Sales are very concentrated in the top markets. The Top 5 countries -- US, Japan, UK, France, and Germany -- represent 76.5% of global sales and

³ This magnitude differs from the one in the introduction because it includes sales of music videos.

the top 10 -- the top 5 above plus Canada, Italy, Spain, Australia, and Mexico -- represent 85% (IFPI, 2002).

Sales are also concentrated in a few companies. The 5 biggest companies -- Universal, Sony, EMI, Warner and BMG -- control more than 70% of the global market of music sales, with the rest of the market share distributed among many independent record labels (indies). These latter labels, in some cases, have an important presence in an individual country, region or continent.

Companies and musicians usually negotiate exclusive multi-year contracts. When producing a new album, artists typically receive an up-front payment and a sales-based payment somewhere between 5% and 13% of the retail price of the record (S&P, 2002).

The compact disc (CD) is the most popular music format, representing 72% of total international units sales. Sales of singles, long plays (LPs) and cassettes (MCs) continue to be replaced by sales of CD albums (IFPI, 2002). Two new formats, DVD Audio and Super Audio CD, are growing but do not yet have an important share. These new formats have higher sound quality (Surround Sound)⁴ and contain some extra content such as video clips, interviews with the artists, etc. The tendency to include more content might have been accelerated by the need to differentiate the product from the illegal online substitute.

There is variability in prices across CDs. While the average price of a CD in the US is US\$14.19 (Businessweek, 2003), 28 of the top 50 albums in Billboard charts costs between US\$17.98 and US\$19.98, and only seven are listed US\$14.98 or less (Billboard, 2002). There is not much information on music prices for other countries. In 2001, the European Commission opened an investigation to study the higher prices in Europe and the divergences in prices inside the European Union.⁵ In 2003, the average price of a CD in the UK is US\$16.80 (British Phonographic Industry, 2003).

Distribution costs of music represent a very important share of total costs. A CD with a suggested price to consumers of US\$16.98 has a price of US\$10.50 to the retailer

4 "Surround Sound" refers to multi-channels of sound -- as opposed to only two channels -- that are located around the listener, both in front and back.

5 It was found that the major labels were including minimum advertised prices linked to cooperative advertising agreements in Germany (Official Journal of the European Union, 2002). A lawsuit alleging the same practices was filed in the US. The case ended with a settlement in which the companies agreed to refund US\$67.4 million to consumers and discontinue their minimum pricing policies (New York Times, 2002).

(S&P, 2002). This latter figure includes distribution costs from the record company to the retailer.

The channels of distribution have been changing. Music stores have been shrinking as a source of sales, and are being replaced by supermarkets, discount stores, department stores and online retail. In the US, music stores' share of sales fell from 62% in 1991 to 42% in 2000 (RIAA). In the UK, supermarkets increased their share of music sales from 11.2% in 1999 to 17.7% in 2001 (IFPI, 2002). Online retail, as a share of total sales, increased from 6% in 2001 to 9% in 2002 in Germany, from 4% to 6% in the UK and remained steady at 3% in the US. Online legitimate delivery became available in 2001 but it is still not an important source of sales.

Digital Music Online

Online delivery is possible in "Motion Picture Expert Group-1 Audio layer 3" format (MP3). This is the same format used in the online sharing of pirate copies. MP3 is a way to compress audio data without significantly compromising sound quality.

Sound recordings are originally represented as waves. When the sounds are digitalized, these waves are sampled many times per second and a file is created. CD quality needs a sampling of 44,100 times per second (44.1KHz). Humans can only hear around 10% of sounds that are recorded on a CD. The MP3 compression system eliminates sounds that are not perceptible to humans and softer sounds when different sounds are playing simultaneously. There are different qualities of MP3 compression (bitrate of the file). Bitrates of 64Kbps up to 192Kbps are standard on the internet, but only files above 160Kbps have quality comparable to CDs.

A 5 minute soundtrack that would take more than 50 megabytes in CD compression format can be reduced to a file of 5 megabytes without affecting the sound quality. A compressed file of this magnitude takes 12 minutes to download with a dial-up internet connection (56Kbps modem downstream), 1 minute and 20 seconds with a regular DSL or cable connection (512Kbps downstream) and 20 seconds with a fast DSL connection (2000Kbps downstream).⁶

⁶ Universities have T3 connections that are much faster (from 3Mbps to 45Mbps downstream).

People can upload (rip) CDs to their PCs hard drives and listen to music on their computer, compress the files to an MP3 format to reduce the storage memory requirement and to facilitate the sharing of the files over the internet, and also convert the files back to a CD format and “burn” CDs that can be played in a regular player.

The development of these technologies could represent a very significant reduction in costs considering that around 50% of music costs are distribution expenses.

There are many alternatives to get music online. They can be divided in two groups: legitimate and illegitimate (under the current law).

Legitimate companies either own the copyright or make deals with copyright owners to distribute their music. Among the biggest companies are Pressplay, owned by Sony and Universal; AOL MusicNet, owned by the other 3 big labels (EMI, AOL Time Warner and BMG) and by RealNetworks; the recently launched iTunes (available to Macintosh users) and the Europe-based OD2, Wippit and Tornado Virtue. Most services offer unlimited “streaming audio”⁷ and “tethered downloads”⁸ for a fixed charge per month of around US\$10 and the possibility to do permanent burnable downloads for around US\$1 a song. Prices of digital tracks in Europe are higher; between US\$1.76 and US\$2.35 a song (Billboard, 2003).

Online legitimate digital delivery has been dominated by piracy. After the shut-down of Napster, illegitimate music online can be found globally on P2P file-sharing services such as KaZaA, Groekster, iMesh, and Gnutella. These services are distributed without charge and allow users to download licensed and unlicensed files, including music, movies, games and software. The amount of music available through these services is higher than on any legitimate site. This is particularly true for some genres. During the last years, there has been a worldwide trend towards domestic repertoire. IFPI attributes this fact, in part, to the wide availability of international repertoire on pirate music services. To check this claim, in section VI, I use changes in domestic and international repertoire shares by country.

⁷ Listen to soundtracks without downloading them to the hard disc.

⁸ It is possible to download a song to the hard disc but it is not possible to burn it into a CD or listen to it in a portable player. Music is no longer available to listen if subscription is ended.

III- The Micro-data

This paper uses a European consumer mail survey by Forrester from October 2001 called Technographics. Forrester is a business research company specializing in the information economy. The fieldwork for the survey was conducted by the market research company Taylor Nelson Sofrés. The sampling methodology is proprietary but is meant to ensure a representative sample in each country. Analogous US data from Forrester have been used intensively in economic literature related to the internet (Goolsbee, 2000 and 2001; Goolsbee and Brown, 2002; Goolsbee and Klenow, 2002).

The survey includes 22,488 observations and is representative of the total 16-and-older population in 7 European countries: France, Germany, Italy, Netherlands, Spain, Sweden and UK.⁹ Three of these countries are among the top 5 music markets and 5 are among the top 10. In 2001, they represented 27.8% of international music sales (IFPI, 2002).

The database contains a discrete $\{0,1\}$ variable of purchases of music during the last month for each respondent. This question is broad in the sense that it is not restricted to off-line purchases. For each respondent, it also contains information about access to the internet; purchases of many goods during the last month including, for example: videos, books, software, event tickets and groceries; ownership of many electronic goods including, for example: walkman, Hi-Fi stereo, cellular phone, DVD player, MP3 player, CD writer and game console, and demographic variables: gender, age, work status, education, household size and household income.

For people using the internet at home, the database contains information on the weekly average number of hours spent online, the type of connection -- DSL, cable, ISDN and dial-up -- and information about the respondent's internet activity, including for example: check email, use search engines, purchase goods online, publish own web pages and download MP3 files.

Table 1 presents summary statistics. Across the overall sample (15,228 observations), 39.3% bought music during the month prior to the interview, 9.0%

⁹ The survey includes 13 European countries. However, music purchases are not available for some countries (Austria, Belgium, Finland, Ireland, Norway and Switzerland). Among the countries considered, 7220 individuals have missing values for at least one of the variables.

regularly download MP3 files and 51.0% have internet access. The percentage of people who bought music is much larger among the group who download MP3 files (55.7%) than among those who do not (37.7%).

Considering only those people who have an internet connection at home (5980 observations), 47.1% bought music during the month prior to the interview and 21.0% regularly download MP3s. Again, the fraction of people who bought music is higher among those who regularly download music (55.0%) than those who do not regularly download music (45.0%).

Table 2 compares internet users and broadband access by country in Forrester's sample with data on the number of internet users by the International Telecommunication Union (ITU). The ITU does not collect the data by itself but it compiles information from different country level surveys.¹⁰ Forrester data shows a higher number of internet users than the ITU. On the other hand, in 2002, the ITU reports a lower broadband penetration at home for the only two countries for which it reports broadband access.

There is heterogeneity in per-capita music sales among the countries considered. According to the IFPI, in 2001, per-capita CD sales per year in the UK were 4.1 units, 3.2 in Sweden, 2.7 in Germany, 2.4 in France, 2.3 in Netherlands, 2.0 in Spain and 0.8 units in Italy. Average per-capita unit sales per year over the countries considered in the sample were 2.47. The micro-data does not contain sales quantities to match to the IFPI data.¹¹ Table 1 also reports probabilities of music purchases by country.

IV- Empirical Approach: Microdata

The goal of this paper is to estimate the causal effect of peer-to peer usage on music sales. It was shown in the last section that a comparison of means indicates that individuals who regularly download MP3s buy more music. This positive relationship persists -- although is not statistically different from zero -- when controlling for many individual level characteristics.

¹⁰ The ITU does not report the date of the year in which the data was collected in each country and the broadness of the measure of internet access -- access at home, overall access, etc --.

¹¹ The micro-data probably include purchases of illegitimate copies of music (commercial piracy).

I consider the following model

$$B_i = b_1 D_i + b_2 X_i + \varepsilon_i$$

where B_i and D_i are discrete $\{0,1\}$ variables indicating the response of having bought music during the month prior to the survey and the response of regularly downloading MP3 files, respectively, X_i is a vector of observed individual characteristics and ε_i represents the error. Column I of Table 3 presents OLS estimates. This regression is done with respondents with and without internet access. Feasible generalized least squares is employed to allow for heteroskedasticity across individuals. Household income was standardized to gross annual euros per year and its logarithm was included in the regression. Age is measured in years. Confidence in English is measured from 1 to 5 with 5 being “very confident”. Education takes values from 1 to 3 for lower, middle and upper. Household size has a cap for 5 or more people in the household. Hours of TV and hours of internet (only for internet users) are the average hours per week and take the values 2, 7, 12, 17, 22, 27 and 30 (with a cap for 30 or more). All the other controls are dummies.¹²

The pirated online product can theoretically be a complement for music in regular formats (CDs, MCs, LPs) as the mean comparison and the OLS regression suggest. Consumers may get to know albums by downloading some tracks from the internet and then deciding weather or not to buy the whole album. Downloading takes time; but this cost could be lower than the cost of getting to know music by going to the store and listening to different albums before purchasing. Also, not having an MP3 player, not being able to burn CDs in order to listen to the pirated copy on a regular stereo system, or some quality difference between the copy and the original may induce the purchase of the album.

However, radios advertise music and there exist technology to burn a good quality CD from an MP3 file in order to listen the pirated copy in a regular player. The presence of unobserved heterogeneity in music tastes across individuals may obstacle the identification of the causal effect. The approach of this paper is to use instrumental

¹² Note that prices are not included in the analysis. However, in this section the coefficient is identified in the cross section, and therefore the controversy about the dynamic of prices might be eluded.

variables techniques. To achieve the goal requires an instrumental variable that is correlated with peer-to-peer systems usage, but is otherwise unrelated to music purchases.

To account for the unobserved heterogeneity, I consider the following model

$$B_i = b_1 D_i + b_2 X_i + \theta_i + v_i$$

where θ_i represents unobserved taste for music and v_i represents the error. If θ_i and D_i are not orthogonal, OLS gives a biased and inconsistent estimation of b_1 . One way to overcome this problem is to find an instrumental variable Z_i that is correlated to D_i but orthogonal to θ_i .

In the database there is information about the type of internet connection and about many uses of the internet. There is also a dummy variable for individuals who read computer magazines. These variables could be employed as instruments. The use of broadband connection as an instrument is explained by the fact that it is considerably faster -- as it was mentioned before -- to download a music file with a high-speed connection. In the cases of uses of internet, there are some uses, as “publish own web pages”, “participate in online auctions”, “ask for technical support online”, “make purchases online” and “make financial transactions online”, that require a higher degree of internet sophistication. The variable “read computer magazines” might also signals a higher degree of sophistication. The idea is that performing these activities is correlated to downloading music but might be otherwise unrelated with an unobserved taste for music.

Column 1 of Table 4 reports the first step regression of “download MP3” on all the potential instruments. “Make purchases online” and “make financial transactions” are not significant. All the other potential instruments are significant.

In column II of Table 3, “download MP3” is instrumented with all the potential instruments that are significant in the first stage regression. The IV coefficient on

“download MP3” is negative and significant. It indicates that downloading MP3s reduces the probability of buying music by 34.9%.¹³

To test for exogeneity, the residuals of the IV estimation are regressed on all instruments -- included and excluded in the second step--. This test of overidentifying restrictions¹⁴ rejects the orthogonality of the instruments to the error term. Probability values are reported in the next to last row of Table 3.

As the exogeneity of the potential instruments is rejected, I instrument only with “have DSL or cable connection” and “publish own web pages”; in the case of this last activity, because it seems to require the highest level of internet sophistication among the activities considered. The estimated effect indicates a reduction of 63.9% in the probability of buying music (Column V of table 3). The test of overidentifying restrictions supports the exclusion of these instruments from the purchase equation. Purchases of music are negatively correlated to age, household size and education and positively correlated to listening to music while online,¹⁵ with being male, with being a student, with the level of confidence in English, with the ownership of a CD writer and a MP3 player and with the ownership of complements of music, such as walkman and Hi-Fi stereo.

The exogeneity test fails when using all the potential instruments but support the exclusion of “DSL or cable connection” and “publish own web page”. This might be indicating that the other potential instruments should be included in the purchase equation. Column VI of Table 3 controls for these variables. The estimated effect is bigger, showing a reduction on the probability of purchasing music of 65.3%.

13 Estimation of models of limited dependent variables (LDV) with endogenous *continuous* regressors has been discussed extensively in the literature. Newey, (1987) proposes an estimator that can be calculated by applying GLS to estimates of the reduced form coefficients including the least square residuals from the first step as an additional regressor. Another estimator, that according to Newey (1987) is less efficient, was proposed by Rivers and Vuong (1988). This estimator is calculated by running a probit, including the exogenous variables, the endogenous regressor and the residuals from the first step as explanatory variables. Newey’s estimator and Rivers and Vuong’s estimator are presented in columns III and IV of Table 3. However, here the endogenous regressor is a dummy variable. See Angrist, (2001) for an analysis of estimation of models of LDV with dummy endogenous regressors.

14 The test is computed multiplying N times R2, where N is the number of observations and R2 is the R2 of a regression of the residuals from the second-stage regression on all the instruments (exogenous variables and instruments not included in the second stage regression). This test is distributed χ^2 with degrees of freedom equal to the number of overidentifying restrictions (number of instruments minus 1).

15 This is not necessarily streaming audio by the context in which the question is asked. The question is “Do you ever do any of the following while online?” One of the options is “Listen to music/radio” and there are many other options, such as “Have the TV on,” “Have friends over,” “Read magazines,” etc.

Given the dramatic change in the estimated coefficient in the instrumented regressions in comparison with the OLS estimate, and even considering that first step regressions and tests of overidentifying restrictions suggest the validity of the instruments, special scrutiny of the results is necessary.

The validity of broadband connection as an instrument may be questioned if people decide to acquire fast connections in order to download music. However, these people would have a strong taste for music and therefore the estimates would understate the negative causal effect and be closer to OLS results. Column VII of Table 3 includes only “publish own pages” as instrument to check for the possibility that “DSL or cable connection”, a possibly invalid instrument, were affecting the results. The coefficient on “download MP3” implies a reduction in the probability of buying music of 74.5%.¹⁶

To have a bias in the opposite direction, the instruments would have to be correlated with a low taste for music. Another possibility is that the instruments are correlated with an unobserved low value of time. Downloading music from an illegitimate site is free of charge. However, there are other costs associated with the process of downloading. An important one is the cost of the time involved in this process. I check that the heterogeneity in the value of time is not correlated with the instruments. This is particularly important for the instruments that try to measure internet sophistication. The risk is that sophisticated internet users are people with a low value of time, and therefore more willing to substitute legitimate music for online pirate music. Using broadband internet connection as instrument may not have this problem because people with high cost of leisure would theoretically choose to have fast connections.

Table 5 presents cross tabulations of the instruments with mean probabilities of purchases and ownership of goods that may be related to a taste for music or entertainment -- buying music, buying event tickets, having a Hi-Fi stereo, having a DVD player, having a game console and a 1 to 10 variable -- with ten for strongly agreeing with the statement -- asking if they are “constantly looking for new ways to entertain myself”. Table 5 also reports a cross tabulation with the number of hours that they watch

16 The correlation between “DSL or cable connection” and “participate in online auctions” is 0.12, between “DSL or cable connection” and “publish own web pages” is 0.15 and between “participate in online auctions” and “publish own web pages” is 0.23.

television, an activity that may be thought to be negatively correlated with the price of leisure, and cross tabulations with two demographics: education and age.

The comparison of means of variables signaling a taste for entertainment shows that, across the overall population, the instruments seem to be positively correlated with a taste for entertainment. The difference in means reduces when considering only internet users. Across the overall population, the number of hours of television watched per week is lower for people who publish their own web pages and people who have fast connections. The difference disappears when considering only internet users. The fact that the impact is still strong when instrumenting only with “have DSL or cable connection” (column VIII of Table 3) is also suggesting that the negative impact is not explained by the possibility of not observing the value of time.

To test for the robustness of the results, columns I and II of Table 6 restrict the regressions to people with internet access at home and with internet users more than three years of online experience, respectively. The regression with individuals with more than three years of internet experience is performed to exclude people who acquire an internet connection after Napster started -- to control for the possibility that people went online in order to download music --. Columns II and III of Table 4 present the first step regressions. The negative effect of piracy remains strong over these groups. The coefficient is not significant when excluding people with more than three years of internet experience. However, excluding people with more than two years of internet experience,¹⁷ the coefficient is very similar to the one shown in Column II of Table 6 (-0.55) and is significant at 5%.

I also investigate if the presence of regional unobservables could be driving the results. This is done by analyzing music purchases of non-downloaders and instrumenting the proportion of downloaders of MP3 files in each region with the regional proportion of people who have DSL or cable connection and the proportion who publish their own web pages. Column III of Table 6 shows that the regional share of downloaders has no effect on the probability of music purchases.

As a final test of robustness, I analyze how purchases of other goods are affected by MP3 downloading; using the same instruments. Column IV of table 6 shows that the

¹⁷ Napster was launched 2 years and 4 month before the fielding of the survey.

effect of MP3 downloading on purchases of videos and DVD is positive and not significant. The swapping of films on the internet has some importance now, but was nearly non-existent in 2001. The sharing of films requires fast connections because the files are big (around 600MB). I have also analyzed purchases of books, software, computer games, event tickets and audio visual electronics, and in none of these cases downloading MP3 has a negative significant effect.

The ownership of an MP3 player may be thought to reduce music purchases by allowing downloaders to listen the pirated files away from the computer. Also, as MP3 files can be converted back to a CD format, “burning” CDs may be expected to reduce music purchases too. The survey does not ask people if they regularly “burn” CDs, but it does ask about the ownership of CD writers.¹⁸ Contrary to the hypotheses, the regressions presented so far show a positive effect of the ownership of MP3 players and CD writers on the probability of buying music.

However, these variables may also be correlated with the unobserved heterogeneity in tastes for music. To analyze the causal effect of the ownership of these goods on music purchases, column V of Table 6 restricts the analysis to music downloaders. As might be expected, the regression shows that having a CD writer reduces the probability of buying music 5.1% and having a MP3 player has a negative (but not statistically different from zero) impact. MP3 players were not popular in 2001 and only 9% of music downloaders (129 individuals) report ownership. On the other hand, 64% of music downloaders report owning a CD writer.

Having a CD writer allows to make copies of CDs even if the music is not pirated on the internet. To test for the effect on sales of people who only burn CDs, I restrict the analysis to people who do not download MP3 files. There are 2098 owners of CD writers who do not download MP3 files. Using the ownership of a scanner and a DVD drive in the computer as instruments, Column VI of Table 6 shows that having a CD writer reduces the probability of buying music in 9.0%. Column IV of Table 4 presents the first step regression.

The above regressions seem to suggest that the effect of CD writers on music purchases is bigger for downloaders than for non-downloaders. Downloading music

¹⁸ It is possible that CD writers come embedded with computers.

online is not legal and some people who actually download music might not want to reveal it. The ownership of a CD writer could be taking the effect of people who actually download MP3 files.

Finally, the last column of Table 6 presents a 2SLS regression using “DSL or cable connection” and “publish own pages” to predict “download MP3”, and “own a scanner” and “own a DVD drive” to predict “Own a CD writer”. An interaction term of both predicted values, to account for people who both regularly download music and have CD writers, is also included as a control. This regression shows that downloading MP3 reduces by 52.1% the probability of purchasing music for people who do not have CD writers and by 53.5% for people who have CD writers. It also shows that having a CD writer reduces the probability of music purchases by 6.3% for non-downloaders and by 7.7% for downloaders. Note that the interaction term is not significant.

V- Effects of online music piracy on music sales

The results in the last section indicate that downloading MP3 files online changes the probability of buying music during the month prior to the survey between 35% and 65%. Pirated music may be shared with people who do not download MP3 and affect their purchases too. In this event, the estimates understate the causal effect of online piracy on music sales. However, instrumenting the proportion of downloaders by region, the results in the last section suggest that there are no indirect displacement of purchases.

The database does not contain quantities of music purchased to calculate what music sales would have been in the absence of online piracy. Another obstacle in calculating the impact on sales comes from the possibility of having downloaders over-sampled or under-sampled (Table 2). Having said that, in this section I investigate this counterfactual.

Across the overall population, 9% regularly downloaded music online in 2001.¹⁹ Therefore, if both digital music users and nonusers had the same propensity to buy music,

¹⁹ Legal online digital delivery was nearly nonexistent in 2001.

the effect on the music industry would be a reduction in music sales units of between 3.1% and 5.8%.

But digital music users have a higher propensity to buy music, indicating that a correction for the heterogeneity in the groups is needed. It was shown in Table 1 that the probability of buying music for the group of people who download music is 50% higher than the probability for nonusers of P2P systems. As there is no data on quantities of music purchased, an assumption about the number of units bought is needed. One conservative assumption is to assume that users and nonusers of P2P systems, reporting that they bought music during the month prior to the survey, bought the same *quantity* of units. With this assumption, the estimated impact on units sold, at the 2001 level of pirate site users, would be between 4.7% and 8.7%. If users of P2P systems not only have higher propensity to buy music, but also buy more units when they buy, this would be an underestimate of the impact.

In year 2002, the IFPI reports a global increase in the number of users from 3 million in 2001 to 5 million (IFPI, January 2003). Assuming that this magnitude is representative of the increase in the European countries, the predicted drop in sales would be between 7.8% and 14.5%.

VI- Aggregate Data: Empirical Analysis with a Panel of Countries

In this section I combine aggregate country level data from different sources. Data on music sales is from the IFPI. The IFPI has a panel of sales for 71 countries for years 1997-2002. Music sales data by country is measured in units, decomposed in singles, long plays (LPs), cassettes (MCs) and compact discs (CDs). For years 1997-2001 and some countries, the panel also contains shares of international, domestic and classical repertoire as percentage of total value of music sales. I combine this panel with data on the number of internet users by country provided by the International Telecommunication Union (ITU), with a panel of software piracy²⁰ by country from the Business Software

²⁰ Software piracy is measured as the percentage of software sales that is lost due to piracy. A demand for new software is estimated using the number of PCs and an estimated ratio of software to PCs in the US market. The supply of software is estimated with shipments data of BSA members companies. Piracy is defined as the difference between software demand and software legally supplied.

Alliance (BSA) and with panels of total and per capita GDP and exchange rates²¹ from the World Bank and IMF, respectively.

The individual level data does not contain quantities of music purchased; only allowing back of the envelope calculations to predict the effect of online piracy on music sales. The turn to the aggregate data implies working with fewer observations and controls, but it allows checking the effect on quantities. Besides, the analysis of aggregate data is another test of the effect of online piracy.

To estimate the effect of the development of online music piracy on music sales, I analyze how the country level of peer-to-peer users affects the change in music sales pre and post existence of P2P systems. Consider a two period model -- pre and post existence of peer-to-peer systems—in which the demand for music in country i and period t is determined by

$$S_{it} = a_t + bP2P_{it} + cX_{it} + e_{it} \quad t = 1,2$$

where S is the logarithm of music sales per capita, $P2P$ is the number of peer-to-peer users per capita and X is a group of other controls. I assume that the relationship between the number of $P2P$ and internet users has the following form

$$P2P_{it} = \begin{cases} 0 & \text{if } t = 1 \\ \lambda I_{it} & \text{if } t = 2 \end{cases}$$

where I is the number of internet users per capita.²² Substituting and taking differences, the change in the logarithm of music sales follows

$$(1) \quad S_{i2} - S_{i1} = (a_2 - a_1) + b\lambda I_{it} + c(X_{i2} - X_{i1}) + (e_{i2} - e_{i1}) \quad t = 1,2$$

Table 7 presents different estimates of equation (1). The dependent variable is the change in the logarithm of the sum of sales in years 2001 and 2002 and the logarithm of

21 Units of foreign currency per dollar.

22 If the proportion of P2P users varies by country ($\lambda_i = \lambda + u_i$), we need u_i to be uncorrelated with I_i .

the sum of sales in years 1997 and 1998.²³ All but the last columns of Table 7 use units of CD as the measure of sales. The covariates are the level of internet users in year 2001, the change in the logarithms of year 2001 and 1997 of GDP per capita, of piracy -- to capture commercial piracy -- and of exchange rates -- to capture changes in prices of music --. Dummies for English being the first spoken language -- English is the dominant language in online sites to swap music -- and for continents are also included.²⁴

Assigning equal weight to all the countries in the sample, the estimated coefficient implies a reduction in music sales of 24.6% at the country mean of per capita internet users.²⁵ Quantile regression, to reduce the weight of outliers, gives a smaller estimated impact, an implied reduction of music sales of 9.5% at the mean. However, there is a lot of heterogeneity of economy sizes among the countries in the sample. To correct for this factor, Table 7 also presents regressions weighted by total number of internet users, by total GDP and by sales of CD units in year 1997. Adjusting for sizes, the implied estimated impact on sales at the mean goes from 13.5% to 16.2%.²⁶

As might be expected, changes in GDP per capita are positively correlated with changes in sales. Also, in the weighted regressions, changes in piracy and in exchange rates show the expected sign, although the effect of a change in the exchange rates is not significant.²⁷

Interacting internet users per capita and the logarithm of total GDP reveals that, given the level of internet usage, an increase in GDP reduces the marginal impact of piracy on sales. For example, for the US, this regression implies a reduction in sales of

23 No population growth is implicitly assumed.

24 The “continents” are Africa, Asia, Australasia, Europe, Latin America, Middle East and North America. “English” is the first spoken language in Australia, Canada, Ireland, Jamaica, New Zealand, UK, USA and Zimbabwe (CIA World Factbook).

25 The mean of internet users per capita within the countries in the sample is 0.206, with a min of 0.003 and a max of 0.599.

26 Without including the US, the coefficient weighting by internet users is -0.5388 , weighting by total GDP is -0.7932 and weighting by sales is -0.6505 , and are all significant at 5%.

27 It was mentioned in section II that, across the countries with information, there is some variation in prices of CDs. If music were a commodity, it is not possible to identify the effect of changes in prices with the proposed model. I also tried including the change in the logarithm of “prices” of music calculated as the quotient of total music sales in dollars and total unit sales (CD, MCs, LPs and Singles). The coefficient on internet users remains unaltered but the coefficient on prices is also not significant.

10.3%.²⁸ For 5 of the European countries included in the microdata the estimated effect is around 20%, and for Netherlands and Sweden the estimated effect is around 50%.

Contrary to what might be expected, English has a positive --though not significant-- impact on sales. The positive effect persists when interacting English with internet users per capita. This result might be suggesting that the level of English proficiency required to swap music online is not too high. It might also be indicating that the availability of repertoire in English online is not relatively more abundant or that music swappers listen to music in the English language. The analysis with the micro-data also showed that the confidence in English does not have a significant effect on the probability of downloading MP3. It should also be mentioned that the measure of English fluency is rather simple. Many countries -- for which English is not the principal language -- have a high understanding of the language.

The restriction to CDs could be misleading if the trend of substitution of cassettes for CDs is correlated with the country level of internet usage. The last two columns of Table 7 explore this possibility including the share of quantities of CDs over the sum of CDs and MCs of year 1997 as a control variable and using the change in the logarithm of sales of MCs as the dependent variable, respectively.

The relationship between growth of MCs sales and internet usage is not significant and the inclusion of the ratio of sales of CD and MC does not change the negative relationship between sales of CDs and level of internet per capita. The non-significant relationship between the growth of sales of MCs and piracy might be suggesting that online piracy is a closer substitute of CDs than of cassettes.²⁹ However, the coefficient is larger in absolute value -- but still not significant -- when weighting by number of users, GDP and sales of MCs.

During the last years, music sales have experienced a trend toward local repertoire. Globally, domestic repertoire increased from 64.0% in 1997 to 67.5% in 2001. According to the IFPI, one reason behind this trend is the wide availability of international repertoire on pirate music services; especially some top-selling international

28 The mean of the logarithm of total GDP for the countries in the sample is 25.32, with a min of 22.54 and a max of 29.93.

29 Column IX of Table 1 presents a regression weighted by sales of units of CD to make it comparable to the regression using units of CDs as the dependent variable.

acts (IFPI, 2002). I use data on domestic and international shares by country to test this claim. In most cases, the classification of domestic-international is defined on the country of signing. In France and the Middle East countries, the classification is defined by language. The US is excluded from the regressions because is the signing place of a high proportion of music acts (US acts represent a high proportion of international repertoire sales in other countries).

I consider how the level of internet users in 2001 affects the change in the logarithms of the shares of international and domestic repertoire of year 2001 and the logarithms of the average shares of years 1997 and 1998 of international and domestic music.³⁰ Table 8 shows results weighted by total number of internet users, by total GDP and by the dollar market value of sales in year 1997.

The results indicate that, as the music industry claims, there is some evidence that online piracy has a stronger effect on international repertoire. However, the impact is non-significant when weighting by total dollar value of sales. Note that data by repertoire is limited to a reduced number of countries. Classical repertoire is available for 38 countries. Restricting to these countries, the effect by repertoire is not significant.

VII- Conclusion

This paper uses two different databases and empirical approaches to measure the causal effect of online music piracy on music sales and finds that file-sharing may reduce the probability to purchase music by around 35% and 65% and may explain an important reduction in music sales. Freely downloadable music seems to be a close substitute of off-line commercially purchased music.

The estimates in this paper are important for welfare analysis. Strong property rights create monopoly distortions but weak property rights may lead to low creation of artistic work. The debate about the underproduction-underutilization trade-off has not been very much informed empirically. Knowledge of the actual estimates of the impact is

³⁰ Note that, as the dependent variable is the change in the logarithm of the shares, the change in the logarithm of the share of one type of repertoire is not equal to the negative of the change in the logarithm of the share of the remaining repertoire.

essential in the determination of the strength of property right. The interest is not exclusive to the music industry. Other digital copyrighted goods, such as movies, software, games, books, etc, are also being swapped online. The development of fast connections will increase the importance of the impact of file sharing on sales of these goods.

Piracy is illegal under the current legal system. The music industry is fighting piracy in court.³¹ In the US, music piracy has been legally fought on the basis of contributory and vicarious liability (Landes and Lichtman, 2003). Under these doctrines, copyright holders sue parties that in some way contribute or benefit from the infringing conduct, instead of suing individuals. However, the new P2P systems are more sophisticated and difficult to fight legally because they do not require a central server to operate (Varian, 2000) and have alternative legitimate uses. The other difficulty is that these new systems are established in countries with different legal systems (KaZaa is registered in the South Pacific island nation of Vanuatu, the software distributor is in Australia and the servers are in the Netherlands; Chicago Tribune, 2003).

Recently there has been a change in the legal strategy. The RIAA has been “gathering evidence and preparing lawsuits against individual computer users who are illegally offering to ‘share’ substantial amounts of copyrighted music over peer-to-peer networks”. In Europe, this strategy is less promising. Recently, the European Commission has proposed to the European Parliament an Enforcement Directive with criminal provisions for infringing P2P file-sharing businesses but excluding not-for-profit copyright infringements (Grammy.com, 2003).

31 In Canada, royalties are imposed on digital media, recorders and players. Supposedly, however, these are royalties to copyright holders for private copying for personal use only and do not compensate for piracy/unauthorized copying of sound recordings online or elsewhere. In Italy and Germany, blank CD and burners pay between 5% and 10%. A legislation to extend taxes to the 15 EU nations has been passed last year and is pending ratification by the national parliaments. USATODAY, (2002). Another way to combat piracy is to flood the net with decoy files. This alternative seems to be under use given that there are many of these files in P2P systems. Lichtman and Jacobson, (2000).

References

Angrist, Joshua, "Estimation Of Limited Dependent Variable Models With Dummy Endogenous Regressors: Simple Strategies For Empirical Practice", Journal of Business & Economic Statistics, 19(1), January 2001.

Billboard, "Declaration Of Independents", June 8, 2002.

Billboard, "Value-Added Tax Poses Problems For EU Digital Sales", June 28, 2003.

Boldrin, Michele and Levine, David, "The Case Against Intellectual Property", American Economic Review, May 2002, 209-212.

British Phonographic Industry, "Market Information", June 2003.

Business Week, "Inside Napster", August 14, 2000.

Business Week, "Big Music's Broken Record", February 13, 2003.

Chicago Tribune, "KaZaA Sues Movie, Recording Industries", January 28, 2003.

CIA World Factbook, available at www.oecd.org/cia.

CNN.com, "Napster Usage Quadruples", September 11, 2000.

CNN.com, "Study: CDs may soon go the way of vinyl", September 3, 2003.

Fine, Michael, "Soundscan Study On Napster Use And Loss of Sales", June 2000, available at www.riaa.com.

Goolsbee, Austan, "In A World Without Borders: The Impact Of Taxes On Internet Commerce", Quarterly Journal of Economics, May 2000, Vol 115(2), 561-576.

Goolsbee, Austan, "Competition In The Computer Industry: Online Versus Retail", Journal of Industrial Economics, 49(4), December 2001, 487-499.

Goolsbee, Austan and Brown, Jeffrey, "Does The Internet Make Markets More Competitive? Evidence From The Life Insurance Industry", Journal of Political Economy, June 2002, 110(3), 481-507.

Goolsbee, Austan and Klenow, Peter, "Evidence On Learning And Network Externalities In The Diffusion Of Home Computers," Journal of Law and Economics, October 2002, Vol XLV (2) (PT. 1), 317-344.

Grammy.com, February 10, 2003.

IFPI, "The Recording Industry In Numbers", October 2002.

IFPI, January 2003, available at <http://www.ifpi.org/sitecontent/press/inthemedi09.html>.

IFPI, "The Recording Industry World Sales", April 2003.

International Telecommunication Union, ITU, available at www.itu.org.

KaZaA.com, "KaZaA Media Desktop Sets Most Downloaded Software Record", May 26, 2003.

Kai-Lung, Hui and Png, Ivan, "Piracy And The Legitimate Demand For Recorded Music", Contributions to Economic Analysis & Policy, Vol 2, Issue 1, Article 11, 2003.

Klein, Benjamin, Lerner, Andres and Murphy, Kevin "The Economics Of Copyright "Fair Use" In A Networked World", American Economic Review, May 2002, 205-208.

Landes, William and Lichtman, Douglas, “Indirect Liability For Copyright Infringement: Napster And Beyond”, Journal of Economics Perspectives, Vol. 17, No. 2, Spring 2003.

Lichtman, Douglas and Jacobson, David, “Anonymity A Double-Edged Sword For Pirates Online”, The Chicago Tribune, April 13, 2000.

Newey, Whitney, “Efficient Estimation Of Limited Dependent Variable Models With Endogenous Explanatory Variables”, Journal of Econometrics, 36 (1987), 231-250.

New York Times, “Suit Settled Over Pricing Of Music CD’s At 3 Chains”, October 1, 2002.

Official Journal of the European Union, September 30, 2002.

Recording Industry Association of America (RIAA), available at www.riaa.com.

Rivers, Douglas and Vuong, Quang, “Limited Information Estimators And Exogeneity Tests For Simultaneous Probit Models”, Journal of Econometrics, 39(1988), 347-366.

Romer, Paul, “When Should We Use Intellectual Property Rights?”, American Economic Review, May 2002, 213-216.

Standard & Poor’s, “Movies And Home Entertainment”, Industry Surveys, November 2002.

The Observer, “Inside The Web Of Thieves”, April 21, 2002.

USATODAY, “Europe’s Music-Piracy Solution: Taxes”, February 6, 2002.

US District Court, Northern District of California, transcript of proceedings, July 26, 2000.

Varian, Hal, "The Internet Carries Profound Implications For Providers Of Information", New York Times, July 27, 2000.

Yahoo Buzz Index, available at <http://buzz.yahoo.com/world/>.

Table 1
Summary Statistics

Overall Population	
% of internet users	51.0%
% of people who regularly download MP3	9.0%
% of people who bought music during the last month	39.3%
% of people who bought music during the last month given that they regularly download MP3	55.7%
% of people who bought music during the last month given that they do not regularly download MP3	37.7%
% of people with DSL or cable internet connection	4.3%
% of people who have their own web pages	3.0%
Number of observations	15228
Internet Connection At Home	
% of people who regularly download MP3	21.0%
% of people who bought music during the last month	47.1%
% of people who bought music during the last month given that they regularly download MP3	55.0%
% of people who bought music during the last month given that they do not regularly download MP3	45.0%
% of people with DSL or cable internet connection	11.1%
% of people who have their own web pages	6.9%
Number of observations	5980
Probability Of Music Purchases	
France	33.0%
Germany	48.5%
Italy	34.0%
Netherlands	30.9%
Spain	27.1%
Sweden	36.5%
UK	53.5%

Table 2
Internet Access

	Forrester (October 2001)			ITU*		
	Access At Home	Overall Access	Broadband At Home	Access 2001	Access 2002	Broadband At Home 2002
France	28.7	39.7	2.0	26.3	31.3	na
Germany	43.7	55.7	4.8	37.3	42.3	na
Italy	40.2	53.1	1.4	26.8	30.1	na
Netherlands	59.7	67.8	15.5	49.0	53.0	19.0
Spain	18.1	33.3	1.8	18.2	19.3	na
Sweden	59.0	67.6	7.7	51.6	57.3	13.0
UK	47.8	59.9	3.9	32.9	40.6	na

* International Telecommunication Union.

Table 3
Overall Population

	I	II	III	IV	V	VI	VII	VIII	
	OLS	IV(1)	IV Probit(1)(a)	IV Probit(1)(b)	IV(2)	IV(2)	IV(3)	IV(4)	
Download MP3	0.0135 (0.0152)	-0.3496* (0.1137)	-0.3177* (0.0622)	-0.3174* (0.0632)	-0.6391* (0.1991)	-0.6533* (0.2278)	-0.7453** (0.3058)	-0.5509** (0.2531)	
Age	-0.0034* (0.0003)	-0.0035* (0.0003)	-0.0048* (0.0004)	-0.0048* (0.0004)	-0.0036* (0.0003)	-0.0036* (0.0003)	-0.0036* (0.0004)	-0.0036* (0.0003)	
Log of income	-0.0055 (0.0068)	-0.0084 (0.0070)	-0.0092 (0.0087)	-0.0092 (0.0089)	-0.0107 (0.0074)	-0.0094 (0.0074)	-0.0115 (0.0078)	-0.0100 (0.0074)	
No internet access	0.0164 (0.0102)	0.0123 (0.0104)	0.0166 (0.0124)	0.0166 (0.0128)	0.0090 (0.0108)	0.0111 (0.0108)	0.0078 (0.0113)	0.0100 (0.0109)	
Listen to music while online	0.0700* (0.0114)	0.1213* (0.0197)	0.1420* (0.0231)	0.1418* (0.0236)	0.1623* (0.0309)	0.1627* (0.0335)	0.1773* (0.0452)	0.1498* (0.0378)	
Household size	-0.0154* (0.0038)	-0.0142* (0.0038)	-0.0162* (0.0046)	-0.0162* (0.0047)	-0.0133* (0.0041)	-0.0133* (0.0041)	-0.0129* (0.0042)	-0.0136* (0.0040)	
Confidence in English	0.0088** (0.0036)	0.0104* (0.0037)	0.0135* (0.0045)	0.0135* (0.0046)	0.0117* (0.0039)	0.0110* (0.0039)	0.0122* (0.0041)	0.0113* (0.0040)	
Male	0.0507* (0.0077)	0.0609* (0.0084)	0.0820* (0.0104)	0.0819* (0.0106)	0.0690* (0.0099)	0.0649* (0.0094)	0.0720* (0.0119)	0.0665* (0.0107)	
Hours of TV	-0.0002 (0.0004)	-0.0001 (0.0004)	-0.0001 (0.0005)	-0.0001 (0.0005)	0.0000 (0.0004)	0.0000 (0.0004)	0.0000 (0.0004)	-0.0000 (0.0004)	
Hours of internet	-0.0025* (0.0008)	0.0020 (0.0016)	0.0021 (0.0018)	0.0021 (0.0019)	0.0053** (0.0026)	0.0053** (0.0026)	0.0070*** (0.0039)	0.0045 (0.0032)	
Student	0.0300*** (0.0180)	0.0608* (0.0209)	0.0592** (0.0242)	0.0592** (0.0245)	0.0854* (0.0259)	0.0879* (0.0279)	0.0944* (0.0328)	0.0779* (0.0288)	
Work full time	0.0168*** (0.0094)	0.0097 (0.0098)	0.0101 (0.0115)	0.0101 (0.0117)	0.0040 (0.0107)	0.0033 (0.0110)	0.0020 (0.0119)	0.0058 (0.0110)	
Education	-0.0066 (0.0052)	-0.0092*** (0.0054)	-0.0086 (0.0065)	-0.0086 (0.0066)	-0.0113** (0.0057)	-0.0117** (0.0058)	-0.0121** (0.0061)	-0.0107*** (0.0057)	
Own walkman	0.0797* (0.0086)	0.0817* (0.0088)	0.0908* (0.0100)	0.0908* (0.0102)	0.0833* (0.0092)	0.0836* (0.0092)	0.0839* (0.0094)	0.0828* (0.0091)	
Own MP3 player	0.0170 (0.0204)	0.0377*** (0.0221)	0.0549** (0.0265)	0.0548** (0.0260)	0.0543** (0.0257)	0.0546** (0.0264)	0.0604** (0.0297)	0.0493*** (0.0266)	
Own Hi-Fi stereo	0.0677* (0.0091)	0.0638* (0.0093)	0.0927* (0.0120)	0.0927* (0.0123)	0.0606* (0.0097)	0.0606* (0.0097)	0.0595* (0.0101)	0.0616* (0.0097)	
Own CD writer	-0.0103 (0.0110)	0.0365** (0.0185)	0.0445** (0.0211)	0.0445** (0.0214)	0.0739* (0.0285)	0.0698** (0.0295)	0.0876** (0.0415)	0.0625*** (0.0348)	
Ask for technical support online	na	na	na	na	na	0.0298	na	na	
	na	na	na	na	na	(0.0226)	na	na	
Read computer magazines	na	na	na	na	na	0.0409*	na	na	
	na	na	na	na	na	(0.0218)	na	na	
Participate in online auctions	na	na	na	na	na	-0.0203	na	na	
	na	na	na	na	na	(0.0282)	na	na	
Dummies	Country- Ownership of goods (5)- Purchases of goods (6).	Country- Ownership of goods (5)- Purchases of goods (6).	Country- Ownership of goods (5)- Purchases of goods (6).	Country- Ownership of goods (5)- Purchases of goods (6).	Country- Ownership of goods (5)- Purchases of goods (6).	Country- Ownership of goods (5)- Purchases of goods (6).	Country- Ownership of goods (5)- Purchases of goods (6).	Country- Ownership of goods (5)- Purchases of goods (6).	Country- Ownership of goods (5)- Purchases of goods (6).
Overidentification test (p-value)	na	0.08	na	na	0.61	0.42	na	na	
Number of observations	15228	15228	15228	15228	15228	15228	15228	15228	

The dependent variable is a dummy variable for buying music during the last month.

(1) Download MP3 instrumented. Excluded (in the second step) instruments: publish own web page, participate in online auctions, ask for technical support online, have DSL or cable and read computer magazines.

(2) Download MP3 Instrumented. Excluded (in the second step) instruments: publish own web page and have DSL or cable.

(3) Download MP3 Instrumented. Excluded (in the second step) instruments: publish own web page.

(4) Download MP3 Instrumented. Excluded (in the second step) instruments: have DSL or cable.

(5) TV, widescreen TV, dish, cable TV, pay TV, set-top box for TV, VCR, PC, digital assistant, handheld game, WAP, camcorder, printer, digital camera, digital camcorder, web camera, game console, DVD player and mobile phone.

(6) Books, videos, video games, software, toys, sports goods, clothing, footwear, jewellery, gifts, flowers, event tickets, electronics, groceries, beer, wine or spirits, tobacco, health and beauty.

Robust standard errors are listed in parentheses. * 1% significance. **5% significance. ***10% significance.

(a) Rivers and Vuoung (1988). (b) Newey (1987)

Table 4
First Step Regressions

	I	II	III	IV
	OLS(1)(a)	OLS(2)(a)	OLS(3)(a)	OLS(4)(b)
Have DSL or cable	0.0830* (0.0171)	0.0720* (0.0175)	0.0794* (0.0254)	na na
Publish own web page	0.0596* (0.0222)	0.0503** (0.0239)	0.0039 (0.0304)	na na
Participate in online auctions	0.0403** (0.0204)	0.0445** (0.0212)	0.0682** (0.0308)	na na
Ask for technical support online	0.0479* (0.0142)	0.0483* (0.0160)	0.0646* (0.0243)	na na
Make financial transactions online	0.0038 (0.0124)	0.0121 (0.0135)	0.0154 (0.0203)	na na
Make purchases online	0.0195 (0.0120)	0.0292** (0.0131)	0.0137 (0.0205)	na na
Read computer magazines	0.0657* (0.0110)	0.0710* (0.0136)	0.0618* (0.0216)	na na
Own a DVD drive	na na	na na	na na	0.2011* (0.0081)
Own a scanner	na na	na na	na na	0.2212* (0.0074)
Age	-0.0003*** (0.0002)	-0.0020* (0.0005)	-0.0023* (0.0007)	-0.0002 (0.0002)
Log of income	-0.0047 (0.0040)	-0.0177*** (0.0101)	-0.0301*** (0.0159)	0.0108** (0.0052)
No internet access	-0.0062 (0.0045)	na na	na na	-0.0207* (0.0076)
Listen to music while online	0.1352* (0.0092)	0.1169* (0.0107)	0.1416* (0.0188)	0.0457* (0.0080)
Household size	0.0034 (0.0023)	0.0090*** (0.0049)	0.0153*** (0.0079)	0.0000 (0.0028)
Confidence in English	0.0028 (0.0020)	0.0017 (0.0046)	-0.0076 (0.0076)	0.0081* (0.0027)
Male	0.0194* (0.0043)	0.0519* (0.0106)	0.0651* (0.0172)	0.0078 (0.0058)
Hours of TV	0.0003 (0.0002)	0.0003 (0.0006)	-0.0003 (0.0009)	-0.0006*** (0.0003)
Hours of internet	0.0096* (0.0008)	0.0087* (0.0009)	0.0096* (0.0013)	0.0070* (0.0005)
Student	0.0882* (0.0131)	0.1011* (0.0225)	0.1205* (0.0355)	-0.0132 (0.0127)
Work full time	-0.0184* (0.0052)	-0.0254** (0.0112)	-0.0265 (0.0185)	-0.0092 (0.0069)
Education	-0.0078* (0.0029)	-0.0132*** (0.0069)	-0.0057 (0.0114)	0.0022 (0.0039)
Own walkman	0.0056 (0.0045)	0.0167 (0.0103)	-0.0067 (0.0170)	0.0133** (0.0062)
Own MP3 player	0.0533* (0.0154)	0.1023* (0.0276)	0.0864** (0.0384)	0.0656* (0.0144)
Own Hi-Fi stereo	-0.0100** (0.0039)	-0.0123 (0.0149)	-0.0117 (0.0241)	0.0228* (0.0074)
Own CD writer	0.1132* (0.0085)	0.1370* (0.0111)	0.1085* (0.0172)	na na
Dummies	Country- Ownership of goods (5)- Purchases of goods (6).	Country- Ownership of goods (5)- Purchases of goods (6).	Country- Ownership of goods (5)- Purchases of goods (6).	Country- Ownership of goods (5)- Purchases of goods (6).
Number of observations	15228	5980	2401	15228

(a) The dependent variable is a dummy variable for regularly download MP3 files. (b) The dependent variable is own a CD writer. (1) Overall sample. (2) Only people who has internet at home. (3) People with internet at home and more than three years going online. (4) People who download MP3 excluded. (5) TV, widescreen TV, satellite dish, cable TV, pay TV, set-top box for TV, VCR, PC, personal digital assistant, handheld game, WAP, camcorder, printer, digital camera, digital camcorder, web camera, game console, DVD player and mobile phone. (6) Books, videos, video games, software, toys, sports goods, clothing, footwear, jewellery, gifts, flowers, event tickets, electronics, groceries, beer, wine or spirits, tobacco, health and beauty. Standard errors are listed in parentheses. * 1% significance. **5% significance. ***10% significance.

Table 5
Difference In Means

	Overall Population		Internet Users	
	Mean	Test For Difference In Means*	Mean	Test For Difference In Means*
Bought music during last month				
Have DSL or cable	0.44	-2.60	0.44	1.58
Do not have DSL or cable	0.39		0.47	
Publish own web page	0.49	-4.21	0.47	-0.17
Do not publish own web page	0.39		0.47	
Have Hi-Fi Stereo				
Have DSL or cable	0.93	-14.23	0.93	-2.31
Do not have DSL or cable	0.79		0.91	
Publish own web page	0.91	-8.61	0.91	-0.13
Do not publish own web page	0.79		0.91	
Have Game Console				
Have DSL or cable	0.35	-4.41	0.34	-0.76
Do not have DSL or cable	0.27		0.35	
Publish own web page	0.37	-4.34	0.34	-1.08
Do not publish own web page	0.27		0.37	
Have DVD Player				
Have DSL or cable	0.20	-7.45	0.20	-4.34
Do not have DSL or cable	0.08		0.12	
Publish own web page	0.20	-6.37	0.21	-4.04
Do not publish own web page	0.08		0.13	
Constantly Looking For New Entertainment				
Have DSL or cable	4.70	-7.99	4.70	-5.02
Do not have DSL or cable	3.89		4.17	
Publish own web page	5.13	-10.51	5.10	-7.31
Do not publish own web page	3.89		4.17	
Bought event tickets during last month				
Have DSL or cable	0.40	-5.54	0.40	-1.69
Do not have DSL or cable	0.29		0.37	
Publish own web page	0.4	-4.72	0.40	-1.10
Do not publish own web page	0.29		0.37	
Hours Of Television Per Week				
Have DSL or cable	15.58	1.59	15.58	-0.99
Do not have DSL or cable	16.11		15.24	
Publish own web page	14.91	3.10	15.02	0.67
Do not publish own web page	16.12		15.30	
Education				
Have DSL or cable	2.17	-8.20	2.17	-0.15
Do not have DSL or cable	1.93		2.16	
Publish own web page	2.41	-14.55	2.42	-7.73
Do not publish own web page	1.92		2.14	
Age				
Have DSL or cable	37.22	14.70	37.22	4.04
Do not have DSL or cable	45.37		39.53	
Publish own web page	34.24	19.20	34.41	8.29
Do not publish own web page	45.35		39.63	

*t test for difference in means with unequal variances.

Table 6

	I	II	III	IV	V	VI	VII
	IV(1)(a)	IV(1)(a)	IV(2)(a)	IV(1)(b)	OLS(a)	IV(3)(a)	IV(4)(a)
Download MP3	-0.7029* (0.2427)	-0.5463 (0.3635)	na na	0.1200 (0.1445)	na na	na na	-0.5218* (0.1623)
Age	-0.0045* (0.0008)	-0.0034* (0.0013)	-0.0034* (0.0004)	-0.0015* (0.0002)	-0.0032** (0.0014)	-0.0035* (0.0003)	-0.0036* (0.0003)
Log of income	-0.0187 (0.0150)	-0.0368 (0.0246)	-0.0067 (0.0060)	-0.0067 (0.0054)	0.0127 (0.0231)	-0.0053 (0.0072)	-0.0081 (0.0069)
No internet access	na na	na na	0.0164 (0.0112)	0.0146*** (0.0080)	na na	0.0125 (0.0106)	0.0045 (0.0114)
Listen to music while online	0.1730* (0.0331)	0.1570* (0.0593)	0.0726* (0.0110)	-0.0175 (0.0223)	0.0593** (0.0291)	0.0756* (0.0127)	0.1542* (0.0246)
Household size	-0.0068 (0.0070)	-0.0177 (0.0115)	-0.0169* (0.0033)	-0.0025 (0.0031)	-0.0059 (0.0124)	-0.0166* (0.0040)	-0.0131* (0.0038)
Confidence in English	0.0183* (0.0070)	0.0058 (0.0111)	0.0087*** (0.0045)	0.0037 (0.0028)	0.0108 (0.0131)	0.0093** (0.0038)	0.0122* (0.0037)
Male	0.1114* (0.0227)	0.1216* (0.0388)	0.0504* (0.0090)	0.0198* (0.0071)	0.0603** (0.0294)	0.0502* (0.0080)	0.0670* (0.0089)
Hours of TV	0.0003 (0.0008)	-0.0012 (0.0013)	0.0001 (0.0004)	0.0012* (0.0003)	-0.0033** (0.0016)	0.0001 (0.0004)	-0.0001 (0.0004)
Hours of internet	0.0064** (0.0029)	0.0057 (0.0047)	-0.0029* (0.0009)	0.0006 (0.0019)	-0.0008 (0.0014)	-0.0024** (0.0010)	0.0054** (0.0021)
Student	0.0732** (0.0370)	0.0474 (0.0576)	0.0349** (0.0165)	-0.0405** (0.0187)	-0.0062 (0.0457)	0.0322 (0.0203)	0.0735* (0.0217)
Work full time	-0.0036 (0.0182)	-0.0062 (0.0286)	0.0183 (0.0130)	-0.0005 (0.0077)	0.0003 (0.0340)	0.0169*** (0.0098)	0.0042 (0.0100)
Education	-0.0112 (0.0104)	-0.0005 (0.0159)	-0.0080 (0.0052)	0.0001 (0.0042)	0.0060 (0.0193)	-0.0076 (0.0054)	-0.0103*** (0.0053)
Own walkman	0.0865* (0.0161)	0.0650* (0.0239)	0.0822* (0.0071)	0.0072 (0.0064)	0.0496 (0.0317)	0.0829* (0.0090)	0.0844* (0.0087)
Own MP3 player	0.0537 (0.0447)	0.0307 (0.0612)	0.0242 (0.0265)	-0.0071 (0.0191)	-0.0193 (0.0433)	0.0314 (0.0233)	0.0593* (0.0227)
Own Hi-Fi stereo	0.0219 (0.0231)	0.0034 (0.0354)	0.0656* (0.0102)	-0.0064 (0.0067)	0.1034** (0.0524)	0.0675* (0.0093)	0.0646* (0.0092)
Own CD writer	0.0967** (0.0403)	0.0566 (0.0512)	-0.0030 (0.0122)	-0.0259 (0.0204)	-0.0515*** (0.0279)	-0.0908** (0.0373)	-0.0632*** (0.0370)
Interaction Download MP3-Own CD Writer	na na	na na	na na	na na	na na	na na	-0.0142 (0.1047)
Regional Share of MP3 downloaders	na na	na na	-0.0241 (0.1478)	na na	na na	na na	na na
Dummies	Country- Ownership of goods (5)- Purchases of goods (6).	Country- Ownership of goods (5)- Purchases of goods (6).	Country- Ownership of goods (5)- Purchases of goods (6).	Country- Ownership of goods (5)- Purchases of goods (6).	Country- Ownership of goods (5)- Purchases of goods (6).	Country- Ownership of goods (5)- Purchases of goods (6).	Country- Ownership of goods (5)- Purchases of goods (6).
Overidentification test (p-value)			na		na		na
Number of observations	5980	2401	13851	15228	1371	13851	15228

(a) The dependent variable is a dummy variable for buying music during the last month. (b) The dependent variable is a dummy variable for buying videos or DVDs during the last month.

(1) Download MP3 instrumented. Excluded (in the second step) instruments: publish own web page and have DSL or cable.

(2) Regional share of downloaders predicted with the regional share of DSL or cable users and the regional share of people who publish their own web pages. SE Clustered by region (52 regions).

(3) Own CD writer instrumented. Excluded (in the second step) instruments: own scanner and own a DVD drive.

(4) Download MP3 predicted with publish own page, participate in auctions and have DSL or cable. Own a CD writer predicted with own a scanner and own a DVD drive.

(5) TV, widescreen TV, dish, cable TV, pay TV, set-top box for TV, VCR, PC, digital assistant, handheld game, WAP, camcorder, printer, digital camera, digital camcorder, web camera, game console, DVD player and mobile phone.

(6) Books, videos, video games, software, toys, sports goods, clothing, footwear, jewellery, gifts, flowers, event tickets, electronics, groceries, beer, wine or spirits, tobacco, health and beauty.

Robust standard errors are listed in parentheses. * 1% significance. **5% significance. ***10% significance.

Table 7
Aggregate Sales

	I	II	III	IV	V	VI	VII	VIII	IX
	OLS(a)	Quantile(a)	OLS(a)(1)	OLS(a)(2)	OLS(a)(3)	OLS(a)(3)	OLS(a)(3)	OLS(a)(3)	OLS(b)(3)
Internet users per capita-2001	-1.1954*	-0.4620**	-0.6587*	-0.7897*	-0.7829*	-7.0471*	-0.8168*	-0.6394**	-0.1525
	(0.4186)	(0.1929)	(0.2416)	(0.2727)	(0.2825)	(1.6308)	(0.2691)	(0.2829)	(1.1747)
Change in the logarithm of GDP per capita 2001-1997	1.8502**	1.8445*	1.3718*	1.4514*	1.8130*	2.2171*	2.1270*	1.8963*	-0.6978
	(0.7950)	(0.3387)	(0.4753)	(0.5260)	(0.5937)	(0.5377)	(0.5781)	(0.5777)	(2.4690)
Change in the logarithm of piracy 2001-1997	0.3259	0.0291	-0.7343*	-0.4133	-0.3680	-0.8326*	-0.3001	-0.3254	2.3070*
	(0.6027)	(0.2648)	(0.2586)	(0.2838)	(0.2317)	(0.2380)	(0.2221)	(0.2258)	(0.9636)
Change in the logarithm of the exchange rate 2001-1997	0.0398	0.0439	-0.0251	-0.0188	-0.0208	-0.0247	-0.0199	-0.0187	0.0455
	(0.0572)	(0.0265)	(0.0207)	(0.0190)	(0.0203)	(0.0181)	(0.0193)	(0.0197)	(0.0847)
English dummy	0.0263	0.2510**	0.1107	0.1411	0.2041**	0.0816	-1.1852**	0.1917**	-1.2008*
	(0.3168)	(0.1117)	(0.1100)	(0.1270)	(0.0837)	(0.0808)	(0.5493)	(0.0814)	(0.3481)
Interaction internet users-logarithm of total GDP	na	na	na	na	na	0.2285*	na	na	na
	na	na	na	na	na	(0.0587)	na	na	na
Interaction internet users-English	na	na	na	na	na	na	4.2826**	na	na
	na	na	na	na	na	na	(1.6755)	na	na
CD units/(CD units+MC units) (year 1997)	na	na	na	na	na	na	na	-0.0019**	na
	na	na	na	na	na	na	na	(0.0009)	na
Other dummies	Continent	Continent	Continent	Continent	Continent	Continent	Continent	Continent	Continent
R2	0.36	0.24	0.53	0.52	0.59	0.68	0.64	0.62	0.68
Number of observations	65	65	65	65	65	65	65	65	65

(a)The dependent variable is the change in logarithms of the sum of CD units sales of years 2001 and 2002 and the logarithm of the sum of sales in years 1997 and 1998. (b) The dependent variable is the change in logarithms of the sum of MC units sales of years 2001 and 2002 and the logarithm of the sum of sales in years 1997 and 1998. (1) Weighted by total number of internet users. (2) Weighted by total GDP 2001. (3) Weighted by CD units sales 1997.

Standard errors in parentheses.*1% significance. **5% significance. ***10% significance.

Table 8
Sales By Origin

	I		II		III	
	International	Domestic	International	Domestic	International	Domestic
Internet users per capita-2001	-0.8841** (0.3430)	0.3850 (0.3081)	-0.9780** (0.4011)	0.2678 (0.3516)	-0.4743 (0.3398)	-0.0692 (0.3425)
Change in the logarithm of GDP per capita 2001-1997	-1.0919 (0.7010)	0.8442 (0.6295)	-1.0058 (0.8022)	0.8222 (0.7033)	-0.2111 (0.7706)	0.7152 (0.7767)
Change in the logarithm of piracy 2001-1997	0.4831 (0.3364)	-0.4940 (0.3021)	0.1541 (0.3779)	-0.0756 (0.3313)	0.0312 (0.2879)	0.1072 (0.2902)
Change in the logarithm of the exchange rate 2001-1997	-0.0006 (0.0241)	-0.0008 (0.0216)	-0.0077 (0.0227)	0.0086 (0.0199)	-0.0181 (0.0226)	0.0177 (0.0227)
English dummy	0.2368 (0.1649)	-0.2561*** (0.1480)	0.2096 (0.1538)	-0.2163 (0.1348)	0.1702*** (0.1004)	-0.1760*** (0.1012)
Other dummies	Continent	Continent	Continent	Continent	Continent	Continent
R2	0.34	0.32	0.26	0.30	0.21	0.29
Number of observations	51	51	51	51	51	51

The dependent variables are the change in the logarithms of the shares of international and domestic music of year 2001 and the logarithms of the averages shares of years 1997 and 1998 of international and domestic music. Shares measured as % of market value.

(I) Weighted by total number of internet users. (II) Weighted by total GDP 2001. (III) Weighted by total dollar value of sales 1997.

Standard errors in parentheses.*1% significance. **5% significance. ***10% significance.