

NBER WORKING PAPER SERIES

PIRACY ON THE HIGH C's:  
MUSIC DOWNLOADING, SALES DISPLACEMENT,  
AND SOCIAL WELFARE IN A SAMPLE OF COLLEGE STUDENTS

Rafael Rob  
Joel Waldfogel

Working Paper 10874  
<http://www.nber.org/papers/w10874>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
October 2004

We are grateful to Betsy Bailey, Mary Benner, Lisa George, Jeff Milyo, and Yochanan Shachmurove for allowing us to administer surveys to their students, and to Hannah and Sarah Waldfogel for help administering the surveys. Lu Chen, David Sokoler and Duygu Baydur provided diligent research assistance. Rob thanks the NSF, and Waldfogel thanks the Wharton Electronic Business Initiative, for financial support. All errors are our own. The views expressed herein are those of the author(s) and not necessarily those of the National Bureau of Economic Research.

© 2004 by Rafael Rob and Joel Waldfogel. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Piracy on the High C's: Music Downloading, Sales Displacement, and Social Welfare in a Sample of College Students

Rafael Rob and Joel Waldfogel  
NBER Working Paper No. 10874

October 2004

JEL No. L82

### **ABSTRACT**

Recording industry revenue has fallen sharply in the last three years, and some — but not all — observers attribute this to file sharing. We collect new data on albums obtained via purchase and downloading, as well as the consumers' valuations of these albums, among a sample of US college students in 2003. We provide new estimates of sales displacement induced by downloading using both OLS and an instrumental variables approach using access to broadband as a source of exogenous variation in downloading. Each album download reduces purchases by about 0.2 in our sample, although possibly much more. Our valuation data allow us to measure the effects of downloading on welfare as well as expenditure in a subsample of Penn undergraduates, and we find that downloading reduces their per capita expenditure (on hit albums released 1999-2003) from \$126 to \$100 but raises per capita consumer welfare by \$70.

Rafael Rob  
Department of Economics  
University of Pennsylvania  
rrob@econ.sas.upenn.edu

Joel Waldfogel  
Public Policy and Management  
University of Pennsylvania, Wharton School  
3100 Steinberg Hall-Dietrich Hall  
Philadelphia, PA 19104-6372  
and NBER  
waldfogj@wharton.upenn.edu

*No black flags with skull and crossbones, no cutlasses, cannons, or daggers identify today's pirates. You can't see them coming; there's no warning shot across your bow. Yet rest assured the pirates are out there because today there is plenty of gold (and platinum and diamonds) to be had. – RIAA website (<http://www.riaa.com/issues/piracy/default.asp>)*

After growing an average of 10 percent per year over the previous 7 years, US music industry revenue has shrunk by 16 percent in the last 3 years (see figure 1). There is considerable debate about causes. Theoretical work on sharing of information products has ambiguous predictions for the effects of file sharing on industry revenue (see Besen, 1986; Bakos, Brynjolffson, and Lichtman, 1999; Varian, 2000; and Shapiro and Varian, 1999); and the empirical literature on the sales-displacing effects of music downloading has achieved no consensus as yet. Two recent studies reach opposite conclusions (see Oberholzer-Gee and Strumpf, 2004, and Zentner, 2003).

One obstacle to empirical consensus is the difficulty of obtaining suitable data, namely information on purchase and download volumes for the same individuals. This study examines the effect of downloading on sales and welfare, asking a series of specific questions. First, using a new survey-based dataset on quantities of music downloading and purchases of 8200 albums by 412 college students at Penn, CCNY, Hunter, and master's students at Chicago's Harris School, we revisit the question of how downloading affects music purchases. Using both OLS as well as an empirical strategy instrumenting downloading volumes with information about broadband access, we ask whether individuals downloading more albums purchase fewer albums. We document that downloading reduces music purchases, by roughly one fifth of a sale for each recent download and possibly much more. Because we have measures of the quantities of music purchased and downloaded, we can estimate the size of the downloading-induced sales displacement, not simply document whether it occurred. Our conservative estimates indicate that downloading reduced recent purchases by individuals in the sample by about 10 percent during 2003.

We then turn to welfare analysis. Two features of the CD market make the welfare analysis of downloading interesting. First, because CDs were easily transferable even in the absence of downloading, substantial price discrimination was impracticable. As a result, firms were compelled to price as single-price monopolists, leaving some socially beneficial transactions (with buyer valuation above marginal cost) unconsummated. This problem was exacerbated by CDs' second feature: the product has a low marginal cost – now that digital distribution is feasible, it is essentially a marketable public good – so the market without unpaid downloading has the potential for substantial deadweight loss.

The advent of unpaid downloading allows consumers to engage in a crude “do-it-yourself” form of third degree price discrimination. Based on access to the Internet and willingness to engage in file sharing, consumers segment themselves into downloaders and buyers. Given the available supply of music, downloaders pay zero and are clearly better off. But that leaves various avenues for downloading to affect welfare more generally. Revenue may decline, harming firms and their suppliers. But revenue will only decline if downloaded albums would otherwise have been purchased. If downloaded albums would not have been purchased, then not only is revenue unaffected but deadweight loss shrinks as well.<sup>1</sup>

We address these issues through surveys asking respondents not just what they bought and downloaded but also how highly they value the music in dollar terms. We document that, on average, our respondents download music that they value a third to a half less than their purchased music. This indicates that at least some of the music that is downloaded would not otherwise have been purchased and therefore supports the incomplete sales displacement that we document.

---

<sup>1</sup>Downloading might also have other interesting effects outside the scope of the current study. First, downloading may change the residual (non-downloader) demand for music in a way causes firms to change prices, harming or helping actual buyers. Second, downloading may change the supply of available music in ways that ultimately change consumers' options.

Moreover, it suggests that some of the surplus enjoyed by downloaders would otherwise have been the deadweight loss associated with a foregone but socially beneficial transaction.

Music has the complicating feature of being an experience good, so that consumers' ex ante valuation (at the time of purchase or downloading) need not be the same as the ex post valuation after experiencing the product. Hence, the ex post valuation, relevant to welfare, deviates from the valuation determining the purchase decision. To address this we administered an additional survey to 92 Penn students eliciting both ex ante and ex post valuation information for 1209 purchased and downloaded music albums. Assuming that albums with ex ante valuations above \$15 would have been purchased, while those with ex ante valuations below \$15 would not, we find that downloading reduces expenditures by individuals in this sub-sample by roughly 20 percent, from \$126 to \$101 per capita. Although downloading reduces revenue by \$25 per capita, it raises consumer surplus by \$70 per capita (for their purchases of hit albums, 1999-2003). The reduction in deadweight loss (\$45 per capita) is nearly double the reduction in industry revenue (from individuals in our sample).

The paper proceeds in four sections. Section 1 provides industry background and links to relevant literatures. Section 2 explores possible effects of downloading on consumption and welfare in theory. Section 3 describes the data used in the study. Section 4 presents empirical results on sales displacement and whether high or low valuation albums are downloaded. Section 4 also presents estimates of the effects of downloading on welfare.

## I. Background

### a. Industry background

The US music industry is highly concentrated. Although there are hundreds of “labels,” five owners account for 80 percent of music sales. The major firms include Sony, Vivendi-Universal, EMI, Bertelsman, and Time-Warner (see Graves, 2004). According to some reports, there are nearly 30,000 CDs released per year. The major labels collectively release about 7000.<sup>2</sup> Sales are very highly concentrated, though. In 2000 the top 250 (100, 50, 10) albums accounted for 48 (32, 22, 8) percent of total sales. Sales in other years are similarly skewed.

After increasing about 10 percent per year 1993-1999, US album sales have declined 16 percent since (see figure 1). The recording industry, represented by the Recording Industry Association of America (RIAA), blames illegal downloading, which they term “piracy” for sales displacement. “Each year, *the industry loses about \$4.2 billion to piracy worldwide -- "we estimate we lose millions of dollars a day to all forms of piracy."*”<sup>3</sup> The decline in demand is substantial enough to encourage consolidation. Sony and Bertelsman have had discussions, and EMI and Time Warner have been rumored to be in discussions (see Helmore, 2003).

Other developments are suggestive of downloading. Internet penetration has increased from 16 percent in 1997 to 47 percent in 2000. And broadband connection, allowing convenient downloading of large files, stood in early 2003 at 40 percent of connected households. Moreover, the sales of a important complementary goods, MP3 players and blank CDs, has increased substantially over the past five years. MP3 player sales grew from 0.5 million yearly in 1999 to an estimated 2.1 million in 2003 (according to CEA Market Research). Of course, one can use such goods for purchased as well as downloaded music, so these developments are merely suggestive of illegal downloading.

---

<sup>2</sup> These figures, reported at <http://www.azoz.com/music/features/0008.html>, are attributed to RIAA.

<sup>3</sup> See <http://www.riaa.com/issues/piracy/default.asp> (accessed January 9, 2004).

Concerned with reductions in revenue – and convinced that music downloading is the cause – the industry has initiated hundreds of lawsuits against ordinary downloaders as a means of deterring downloading activity. There is some evidence that the lawsuits are having an effect (Schwartz, 2004). In concert, Universal reduced its list prices by 30 percent in October 2003 (USA Today, 2003).

The music industry, initially slow to offer sales over the web, has moved recently to make music available online. They were preceded by Apple Computer, maker of the leading MP3 player, who launched their iTunes music site in 2002. There consumers can download music for \$0.99 per song. As of March 2004, customers had downloaded 50 million songs from iTunes (see Apple Computer, 2004).

#### b. Relevant Literature

This study is related to three strands of academic research. First, our general question, of how differentiated product firms' ability to harvest revenue from potential buyers limits the types of products the firm can market, relates to the product selection problem discussed by Spence (1976) and Dixit and Stiglitz (1977). In general, the structure of demand suggests what set of products would optimally be offered. Which products are actually offered in a market depends on how well firms can appropriate consumers' valuation as revenue whether firms can appropriate. Of course, downloading substantially undermines the ability of firms to appropriate valuations as revenue.

Closer to the phenomenon, our study is related to existing research on the effect of reproduction technology on markets for information goods, beginning with Liebowitz (1985). Liebowitz documents that the advent of widespread photocopying induced academic journal

publishers to raise their prices to libraries as they lost much of the individual market to photocopiers. The issues addressed in this study also echo the issues in the software piracy literature.<sup>4</sup>

This paper is also part of a literature on the effect of file sharing on the revenue from selling information. A number of theoretical papers make the point that file-sharing need not reduce seller revenue (see Bakos, Brynjolffson, and Lichtman, 1999; and Varian, 2000). To see this, imagine two individuals each valuing a good below its price. If they can share it, they may together be willing to purchase the good that they alone were not willing to buy. Shapiro and Varian (1999) also emphasize possible sales-stimulating effects of free samples. Thus, for a number of reasons, whether file sharing displaces sales is an empirical question.

Finally, this is one of a number of recent studies of the effect of music downloading on album sales. Liebowitz (2003) examines a variety of possible explanations for the recent reductions in album sales and, finding them all wanting, concludes that downloading must be responsible. Zentner (2003) uses international time series aggregate data, in conjunction with Internet connectedness, to document that places with more Internet – and broadband – connections have experienced sharper reductions in album sales. He also uses micro data to show that persons who self-report downloading music – instrumented with measures of technical sophistication – are also less likely to have purchased music recently. Hui and Png (2003), using international panel data, 1994-1998, estimate that each download reduces sales by 0.42. The time period they study predates the growth of broadband and widespread file sharing. Moreover, their measure of piracy has the shortcoming that it is an estimate based in part on the level of legitimate sales in a country.

On the other hand, Oberholzer-Gee and Strumpf (OS, 2004) examine weekly time series data on downloading and sales of major hit albums, finding little relationship. And in expert

---

<sup>4</sup> See Besen (1986), Conner and Rumelt (1991), Takeyama (1997).



testimony on the Napster case, Fader (2000) argues that file sharing stimulates sales of recorded music, although his analysis covers a period prior to the sales decline depicted in figure 1.

It is difficult to obtain data well suited to measuring the possible sales-displacing effects of downloading. Album sales data exist, by album and by time. OS have combined weekly sales data with novel data on the weekly volume of downloads to ask whether albums are purchased less when in the weeks when they are downloaded more heavily. They identify sales displacement using within-album weekly variation in downloading and sales. In effect, they ask whether an album sells fewer copies in a week that it is downloaded frequently. Given the durable nature of music files, it is not clear that the absence of contemporaneous substitution rules out substitution more generally.

Their *contemporaneous album sales approach* faces the basic handicap that the variation in a particular album's popularity over time would tend to induce a positive relationship between purchases and downloads. Purchase and download are simply two ways of obtaining an album. To see this clearly, consider a different pair of channels for obtaining an album, from Sam Goody or from HMV. Virtually no one purchases a particular album from both outlets. Thus, the displacement is one-for-one. Yet, the weekly sales of a particular album at Sam Goody are surely highly positively correlated the weekly sales of the same album at HMV. Do HMV sales stimulate Sam Goody sales? Probably not. Downloading and purchases have the same problem. An album's popularity has a time component to it; and when it is popular through one distribution channel, it is popular through other channels as well. And this does not mean that availability through one of the channels stimulates demand through other channels.

It is not even necessary for the relationship between album sales and downloads to inform the sales displacement question. Suppose there are two types of people, buyers and downloaders.

Suppose that buyers never download, and downloaders never buy. Then a negative correlation between sales and downloads across albums just means that buyers and downloaders have different tastes. What we really want to know is whether a *person's* downloading of an album reduces the probability that *he* will purchase the album. Hence, it is not enough to have sales and downloads by album. Rather, one needs data on sales and downloads by *person*, or groups of persons.

The ideal data for studying sales displacement would be volumes of sales and downloads, by individual rather than by album. If one could find exogenous variation in downloading across individuals, then displacement could be inferred from the relationship between downloads and sales (i.e. “do people who download more albums, or songs, purchase fewer albums?”). To our knowledge, surveys are the only way to obtain this information. Zentner (2003) has made use of European individual-level survey data on music downloading and purchases, as well as a host of individual level characteristics. Unfortunately, Zentner’s micro data include only binary measures of downloading and music purchase (do you download music? Have you purchased music recently?), preventing him from measuring the size of downloading-induced sales displacement.

A lack of alternative data sources led us to undertake our own surveys. The benefit of this data collection is that it allowed us to get information on individuals’ volumes of music downloads and purchases. Because we surveyed the college student populations available to us, our surveys are necessarily not nationally representative. We cannot therefore draw inferences about the magnitude of aggregate downloading on US album sales. But we can ask whether the sales displacement phenomenon operates at all. That is, do people who download more purchase less? If so – and if we believe we have isolated exogenous variation in the volume of downloading – then we can at least draw inference about whether downloading – again for our sample – tends to stimulate or cannibalize album sales. This is the spirit in which we proceed.

## II. Theory

The possible effects of illegal downloading on welfare are best analyzed in two parts. First, we discuss the analysis when there is no distinction between ex ante and ex post valuation with simple demand and supply examples. This conveys much useful intuition but misses some important issues that arise with experience goods. We then discuss how the analysis must be adapted for experience goods. Throughout, we ignore possible effects on supply.

### a. Effects of Downloading on Sales and Welfare

Prior to the advent of unpaid downloading but after digital distribution is technically feasible, we view the seller of each album as a single-price monopolist in a context with essentially zero marginal costs. For most demand structures, this will lead to deadweight loss, as sellers restrict output to maximize profit. See figure 2, panel 1, which assumes zero marginal cost.

Downloading segments demand into two groups. Prior to downloading there is some overall demand function for music CDs,  $D(p)$  reflecting the distribution of consumers' willingness to pay for music (each element along the function is a consumer-album). The area under the overall demand curve provides a welfare baseline: with zero marginal costs, this area is the sum of revenue, consumer surplus, and deadweight loss. After the advent of downloading, consumers divide themselves into buyers and downloaders, with aggregate valuation functions,  $D_B(p)$  and  $D_D(p)$ , where at any price,  $D_D(p) + D_B(p) = D(p)$ . Note, however, that for downloaders the valuation function simply measures only hypothetical willingness to pay if downloading were not feasible.

We can work through some simple illustrative examples, although a more realistic analysis require information on what  $D(p)$  and  $D_B(p)$  actually look like, a topic we address empirically

below. Still, the examples help to fix ideas. First, consider a case with linear demand and in which downloaders are drawn uniformly from the overall demand distribution. Then the buyer demand curve is simply an inward rotation of  $D(p)$ . One can easily show that downloading then raises consumer surplus, while reducing revenue and consumer surplus. Note that with linear demand the monopolist's profit-maximizing price is invariant with rotations of the linear demand curve.

Two other polar cases, illustrated in the remainder of figure 2, differing in whether highest or lowest value users download, are instructive. First, consider a case in which low-value users download, in which the pre-downloading price is  $p$ . When downloading is feasible, all users with valuations above  $p$  continue to purchase, while all consumers with valuations between  $p$  and 0, which is the marginal cost of downloading, download. In this extreme case, revenue, and the consumer surplus of persons who bought prior to downloading, are unaffected by downloading. The deadweight loss existing prior to downloading is transformed into consumer surplus. See panel 2.

At another extreme, the high value consumers – say, those with valuations above  $p'$  – could download. For example, if  $p' = p$ , then regions representing CS and revenue prior to downloading become downloader surplus. The region formerly representing DWL now contains revenue, additional CS as well as some deadweight loss if the price is above MC (as it would be for a profit-maximizer). See panel 3.

While these special cases based on simple – and restrictive – representations of demand are suggestive, realistic analysis of the welfare effects of downloading require information about the actual distribution of valuations for buyers and downloaders. The key question, for both the welfare analysis of downloading and the size of the sales displacement, is whether individuals download high or low-valuation albums. The more that consumers obtain low-valuation music

through downloading, the smaller its negative impact on revenue and the greater its beneficial effect on welfare.

#### b. Decoupling Ex Ante and Ex Post Valuation

Because music is an experience good, the ex ante valuation determining purchase is decoupled from the ex post valuation relevant to welfare. As a result, actual welfare is slightly more complicated than the foregoing section would suggest. At the time of first obtaining the album, the individual's willingness to pay is based only on a guess about how much she will like an album. Call this the ex ante valuation,  $v_i^a$ , where "a" denotes ex ante, and (i) indexes a buyer-album pair. Her ex post valuation of the album, once she has listened to it for a while, can deviate from  $v_i^a$ . Term the ex post valuation  $v_i^p$ , where the "p" denotes ex post. Define  $p_i$  as the purchase price the buyer faces. Finally, each consumer-album pair has a binary willingness to download ( $d_i=1$  if yes, 0 if no).

There are two regimes, a downloading regime (when downloading without paying is feasible) and a non-downloading regime (when it is not). When downloading is feasible, each consumer-album pair requires first a decision about willingness to download. If  $d_i=1$  (yes), then downloading occurs as long as  $v_i^a > 0$  (regardless of whether the ex ante valuation exceeds the price). If  $d_i=0$ , then she purchases only if  $v_i^a = p_i$ . If  $d_i=0$  and  $v_i^a < p_i$ , she goes without, even though she has valuation in excess of downloading's zero marginal cost.

In the non-downloading regime, individuals purchase only when  $v_i^a = p_i$ , even though they have valuations in excess of marginal cost. To make matters simple, assume that what is infeasible in the non-downloading regime is illegal downloading but that legal – and revenue-generating – forms of zero-marginal cost distribution are available.

While individuals decide whether to obtain albums on the basis of  $(v_i^a, d_i, p_i)$ , their ultimate satisfaction from the album depends on  $v_i^p$ , less a price paid, if any. Hence,  $CS = v_i^p$  for downloaded albums, and it equals  $v_i^p - p_i$  for purchased albums. Table 1 summarizes the circumstances in which people buy and download in the downloading and non-downloading regimes by partitioning space according to whether  $d_i = 1$ , and whether  $v_i^a = p_i$ .

One point bears discussion here. Because our survey – discussed below – elicits information only about albums that individuals have in their possession (either via purchase or downloading), we will have data on  $v^a$ ,  $v^p$ , and  $d$  for only three of the four cells in the tables. We have no valuation information about albums that people are willing neither to download nor to purchase (with  $d = 0$  and  $v_i^a < p_i$ ). Note, however, that the deadweight loss in this cell is identical across downloading and non-downloading regimes, so that our inability to measure this will not inhibit our ability to measure the change in welfare from downloading.

### III. Data

The basic data for this study are derived from two surveys administered to roughly 500 college students (at Penn, Hunter, Chicago's MA program in public policy, and CCNY) between December 2003 and February 2004. The first survey, administered to 412 students, elicits information on the respondent (age, race, gender, family income number of CDs owned, speed of Internet access, as well as interest in music), along with ex post valuation on two groups of albums purchased or downloaded: those obtained in the last year, and those among a common list of 261 hit albums released since 1999.

Table 2 presents some characteristics of the individuals in the survey one sample. Blacks make up 9 percent of the sample, Hispanics 10 percent, Asians 32. The remainder are white. Fourteen percent of the sample has family income below \$25 thousand; 20 percent has family income between \$25 and \$50 thousand; 28 percent has family income between \$50 and \$100 thousand; 22 percent has family income between \$100 and \$250 thousand, and the remaining 16 percent has family income in excess of \$250 thousand. The mean age is 21.9.

The respondents report high interest in music. Only 15 percent report being less of a music fan than others they know. Nearly forty percent claim to be about the same, 30 percent are somewhat more, and 17 percent are a lot more interested in music than others they know. The mean CD collection is 103.

In a second wave of the first survey we asked about current and past Internet access. We obtained this information on 260 of the 412 respondents. Seventy percent report having dialup access in 1999, while 21 percent had broadband (DSL, cable modem, or Ethernet) access in 1999. By 2003, roughly 85 percent had high speed access, while the remainder had dialup access.

Respondent characteristics vary substantially across the universities sampled. Chicago and Penn are heavily white and Asian, while Hunter and especially CCNY have larger black and Hispanic populations. Chicago income is low because the MA students tend to report their own, as opposed to parents', income (note their comparatively high ages). Penn income is substantially higher than elsewhere. Reported interest in music is somewhat higher at Penn and Hunter than elsewhere. Finally, patterns of Internet access differ substantially across school. Penn students, most living in wired dorms, have virtually 100 percent broadband access. Rates elsewhere are much lower.

The second part of survey one asks respondents to attach dollar value to two groups of albums they have. First, we present them with a list of all 261 albums certified by the RIAA as having sold 2 million or more copies since 1999. For each album, respondents indicate whether they own it, how they got it, and how highly they value it in dollar terms. We term this the “hit” module, and responses to this module provide valuations on a common set of albums across all respondents. Second, we ask respondents to list all albums they have obtained in the past year. For each album they list, we ask them how they obtained it (via purchase, downloading, or as a gift) as well as their valuation. We term this the “current” module.

Two different versions of the valuation question were used initially. The “buy” set asked respondents, “Imagine you no longer have the album and must pay to get it. What is the maximum you are willing to pay to get the album.” In the alternative “sell” treatment, the question is worded, “Imagine someone offered to pay you money to permanently give up the album (and never again hear any of the songs on it). For each album, how much compensation would you require never to hear it again?” We employed the buy formulation for 337 of 412 surveys.

We use the survey data in two ways. First, we can aggregate them to the respondent level to create a cross-respondent dataset on the volume of CDs purchased and downloaded. Although we do not know when respondents purchased each CD, we do know the CDs’ “vintage,” i.e. the year in which they were first certified. This allows us to create a second, panel dataset on the numbers of albums downloaded and purchased, among those of each vintage. Third, we can use the data at the album level.

Table 3 reports mean and median valuations of purchased (including gifts) and downloaded hit and current albums using both the buy and sell valuation methods. The first point to note is that sell-based valuations are enormous. The average purchased hit is valued at over \$50,000. While



the means are sensitive to outliers, even the median sell valuations - \$100 per album for hits and current albums – are rather high. The buy-based valuations appear to be more reasonable. Of 2726 purchased hits, the mean valuation is \$12.70, while the mean valuation of a purchased current album is \$15.25. Second, downloaded albums are valued substantially below purchased albums.

Although more reasonable in the sense of being within an order of magnitude of purchase prices, the survey 1 buy valuation data are still, at first blush, puzzlingly low. Of 2726 purchased albums valued via the buy method, the mean valuation was \$12.70. On its face this seems to contradict logic, since a person should only buy a product if the value equals or exceeds the price. And, indeed, 58 percent were valued below \$15. Moreover, 28 percent were valued below \$10. These results raised concerns for us that music may be a) an experience good, b) subject to depreciation as listeners grow tired of music, or c) both.

Low ex post valuations are not just a theoretical curiosity; they also prevent using survey 1 valuation data to determine which downloaded albums would have been purchased absent downloading. It seems natural to assume that individuals would have purchased albums when  $v > p$ . Yet, it is not true that  $v$  exceeds  $p$  for all, or even most, purchased albums.

Our concern about music as an experience good led us to administer survey two to 92 Penn students, asking about both ex ante and ex post valuations of the 261 albums in the hit sample, as well as the individual characteristics covered in survey one. Our valuation instructions on this survey were rather explicitly designed to elicit ex ante valuations at or above the prices paid:

**Initial valuation** - *At the time you obtained the album, what is the maximum you would have been willing to pay for it? If you paid for it, you must have valued it at least as much as its price initially. For this question, suppose that there is only one possible source. That is, your answer should not be based on an alternative price or sharing opportunity. (Don't say you're not willing to pay since you know you can obtain it via sharing for free, or that you are not willing to pay more than some particular amount since you know you can buy from another seller for some particular price. Here, you are asked to assume no such alternative sources exist.)*

Finally, if you would have been willing to purchase it at the time you downloaded it or received it as a gift, report a valuation at or above the going price at that time.

**Current valuation** - *Now that you've had the music for a while, what is the maximum you would be willing to pay for it? Here, again, suppose that there is only one possible source so your answer, as above, should not be based on an alternative price or sharing opportunity.*

The last column of table 2 reports characteristics of the individuals in the survey 2 sample. The bottom panel of table 3 reports ex ante and ex post valuations of hits from survey 2. Here, the mean ex ante valuation is nearly \$16, while the mean ex post valuation, at \$13.39, is very similar to the (ex post) valuation in survey 1. Two points merit discussion. First, the ex post valuation is below the ex ante, indicating depreciation. Second, the correlation of ex ante and ex post value is only about 0.6, indicating that most of the variation in ex post value is realized after purchase, so the experience good aspect of music is important.

We supplement the survey data with information on aggregate sales data on all of our hits – as well as roughly 8000 other hit albums certified by RIAA – from the Recording Industry Association of America Website (RIAA.COM). The source reports individual certifications. An album is certified “gold” when it has sold 0.5 million units, “platinum” when it sells 1 million, and it receives a multiple platinum certification for each additional million copies sold. An album released in, say, 1985, can continue to sell in subsequent years. We attribute the additional sales leading to most recent certification to the year of the current certification. This makes the certification-based data compatible with the RIAA aggregate sales data (except that the aggregate data also include albums selling fewer than 0.5 million copies). RIAA also provides data on annual units sold and revenue. Figure 1 shows the overall and certified album sales. RIAA certified albums account for roughly three quarters of overall sales.

Table 4 lists the top 40 albums in the hit sample, by number of appearances among the hit albums in our respondents' collections. The last column reports the certified sales of each album (as of late 2003). Although we make no claims of representativeness in our sampling, the albums

with high sales tend to appear frequently in our sample. For example, the Backstreet Boys and 'N Sync albums which each sold over 10 million copies each appear among more than 50 of our respondents' hit album collections.

Table 5 presents the average number of purchases and downloads, in the current and hit samples, overall and by race and school.<sup>5</sup> Respondents report an average of 7.8 hit purchases and 5.1 hit downloads (for albums first certified in 1999 and by late 2003 selling at least 2 million copies). Blacks download fewer albums than others in the sample. Penn students in our sample download substantially more albums than Hunter, CCNY, or Chicago students (7.2 hit albums as opposed to 2-3). Interestingly, Penn students purchase fewer albums than the others, suggesting sales displacement.

We can compare our self-reported purchase information with aggregate data on album sales. According to the RIAA, US album sales in 2002 totaled \$11.233 billion. Of this total, 11.5 percent, or \$1.292 billion, was sold to persons age 20-24. According to the 2000 Census, there were 18.964 million persons in this age group in 2000. Dividing, we get average annual expenditure of \$68 per person aged 20-24. Given the price of CDs this translates to just under 5 albums per person. Our respondents, by contrast, report purchasing an average of 2.6 albums in the past year.

Table 6 describes music obtained by the respondents – and percent downloaded – for hit albums in our list, by date of first certification, 1999-2003. The number obtained decline over time because albums released later have fewer years of sales exposure. Roughly a third of hit albums released in 1999 are obtained by downloading. The share rises to nearly half for albums released

---

<sup>5</sup> In this paper we group albums received as gifts with purchases. All exercises reported in the paper were also performed excluding gifts from purchases. No substantive results change.

during 2002-3. Relative reliance on downloading varies substantially across some groups in our sample, with Penn students downloading proportionally more than others in the sample.

#### IV. Sales Displacement

##### a. Does downloading reduce album sales?

Aggregating our survey data to the respondent level creates data on album purchases and downloads for 412 individuals. Of these, 364 have valid data on all variables needed for basic analysis of the hits. Our first pass at the sales displacement question is through cross sectional regressions of the number of albums purchased on the number of albums downloaded and a host of controls. In particular, we estimate the following equation:

$$P_i = X_i \mathbf{b} + \alpha D_i + \mathbf{e}_i, \quad (1)$$

where:  $P_i$  is individual  $i$ 's purchases of albums from some set (either the hit or current),

$D_i$  is individual  $i$ 's downloads of albums (again from some set),

$X_i$  includes characteristics of the individual such as their level of interest in music, race, and income,

$\mathbf{e}_i$  is unobserved characteristics, and

$\alpha$  and  $\mathbf{b}$  are coefficients to be estimated.

Columns (1-3) of table 7 reports estimates of equation (1) using the entire hit album sample. In this equation,  $P$  is the total number of albums on the hit list (those albums selling 2 million or more between 1999 and 2003) that the individual has purchased, while  $D$  is the number of those albums that the person has obtained via downloading. The first column shows that the estimated sales displacement coefficient  $\alpha$  using the full time period is insignificantly different from zero. The result does not change if we control for school (column 2) or estimate a tobit rather than an

OLS model (column 3). Columns (4-6) present estimates of equation (1) using the current sample. The OLS regressions in columns (4) and (5) give displacement estimates of  $-0.17$  to  $-0.19$ , changing little when we control for school. Column (6) reports a tobit estimate on the current sample, yielding a negative and significant displacement estimate.<sup>6</sup>

In interpreting these estimates, it is important to keep in mind that broadband Internet access stood at low levels for our respondents in 1999. If a person is not connected, there is no scope for downloading to displace her purchases. Many, if not most, of the persons in the sample purchased 1999 releases prior to having the ability to download. The sales displacement question is addressed cleanly only when people have the ability to download albums at the time they might otherwise have purchased them. These considerations suggest that recent years would be most useful for detecting possible displacement and explain why we see displacement in the current estimates but not for the 1999-2003 hits overall. When we estimate equation (1) only on the hits first certified in 2003, we find displacement of  $-0.09$  ( $p\text{-val} = 0.10$ , two-sided).

The estimates in table 7 are vulnerable to the concern that downloading is endogenous, or at least that it is subject to unobserved heterogeneity. We might expect music lovers to both download and purchase more music. One way to deal with this is simply to recognize that it would lead OLS to underestimate the amount of sales displacement and to view our coefficients as underestimates of true sales displacement. Alternatively, one can try to find instruments for music downloading that are not themselves related to music demand. The speed of one's access to the Internet can serve as such an instrument, provided that connection speed is exogenous to interest in music. It would undermine this strategy, for example, if people chose high-speed connections because of their interest in music.

---

<sup>6</sup> Because purchases are positive for 243 of 347 observations, the associated displacement estimate is  $-0.18$ .

We have two approaches to using access speed as an instrument for downloading. First, we can use the school the student attends; and second, for a subsample, we can use whether she has broadband access. Far more than college students at Hunter or CCNY, or MA students at Chicago, Penn students live in dormitories with high-speed Internet access. “Only 612 of Hunter's 17,000 students have the opportunity to live in the Residence Hall” ( see <http://www.hunter.cuny.edu/~reslife/index.html>). And there are no dormitories at CCNY (see [http://www.ccny.cuny.edu/bulletin02/about\\_usU.htm](http://www.ccny.cuny.edu/bulletin02/about_usU.htm)). By contrast, there is space for 5683 of the “nearly 10,000 undergraduates” at Penn. Our own survey shows this: broadband access is ubiquitous among Penn students, far less common among others. Our two IV strategies are (1) to use which school they attend to instrument downloading, and (2) to use the survey’s individual measure of broadband access as an instruments.

The first four columns of table 8 present first and second-stage regressions on the hit sample. The first two columns report first-stage (download) regressions, using school attended and the individual broadband access variables, respectively, as instruments for downloading. As one would expect from the raw data in the tables above, Penn students download more than other students, after accounting for race, income, and interest in music. Columns (3-4) report second-stage hit displacement regressions, and the estimated hit sales displacement coefficients are  $-0.25$  and  $-0.57$ . Columns (5-8) repeat the exercise using the current sample, and again, individuals with broadband access download more, and the estimated current sales displacement coefficients are  $-0.88$  and  $-1.47$ . The estimated IV displacement coefficients in table 8 are much larger than their OLS analogues in table 7, although they are also less precisely estimated.

The panel structure of our data allows another empirical approach. Table 9 uses the panel data, in which a unit of observation is a person-by-vintage for the hit sample. We estimate equations of the form:

$$P_{it} = X_i \mathbf{b} + \alpha D_{it} + \mathbf{f}_t + \mathbf{e}_{it}.$$

Thus we run regressions of the number of albums released in each year that an individual purchased on the number of that year's albums that the person downloads, along with time dummies ( $\mathbf{f}_t$ ) and our various controls. This pooled cross-section time series estimates give sales displacement coefficients of roughly  $-0.04$  but not significant. The second column of table 9 incorporates individual fixed effects, identifying the sales displacement coefficient from the relationship between changes in the tendency to download and changes in the tendency to purchase, over and above the common vintage pattern. That is, we make the substitution  $\mathbf{e}_{it} = \mathbf{m}_i + \mathbf{n}_{it}$ , where  $\mathbf{m}_i$  is an individual-specific fixed effect, and  $\mathbf{n}_{it}$  is an individual-and-year-specific error. Here we obtain a coefficient estimate of  $-0.08$ , and it is statistically significant. When we restrict attention to persons purchasing more than 5 albums from the hit sample we get a coefficient twice as large ( $-0.15$  and significant).

What do all of these estimates mean? Estimates based on data for periods when downloading is feasible suggest sales displacement of roughly  $-0.2$  per album downloaded in the OLS specifications and much higher in the IV estimates. Average current year downloads – see table 4 – were 1.24, while current purchases averaged 2.61. Using the  $-0.2$  displacement estimate, in the absence of downloading, purchases would have averaged 2.86, suggesting that downloading reduced purchases by individuals in the sample by about 9 percent.

- b. Do consumers download high or low value music?

Before analyzing the valuation data, it is worthwhile to discuss the virtues and shortcomings of survey-based valuation information. Economists have traditionally been skeptical of surveys of valuation because a respondent's valuation of an object he has never thought about – or never will see – may not have any meaning (Diamond and Hausman, 1994). In our context, however, we are asking about familiar objects. Indeed, we are in particular asking about objects that the respondents actually possess. Moreover, these are objects of a sort the respondents have obtained repeatedly, so their valuation responses may be more meaningful than responses about unfamiliar and largely hypothetical goods.

Hence, we are willing to put aside qualms about survey valuation data to ask whether downloading is a means of obtaining high or low valuation albums. To this end, Table 10 reports album-level regressions of log valuation on a downloading dummy. The top panel reports regressions of log (ex post) valuation on a downloading dummy, using data from survey one. Only buy valuation data are included in these regressions (to ensure comparability with survey two results). The first column shows that downloaded hits are valued 33 percent below purchased hits. The second column shows that downloaded current albums are valued 39 percent less than purchased current albums. Standard errors in all table 10 regressions are adjusted for clustering on individuals. Columns (3) and (4) repeat the exercise with individual fixed effects, showing not only that downloaded albums are less valuable than purchased albums but also that this is true within individual.

The bottom panel of table 10 performs a similar exercise on the ex ante and ex post hit valuation information from sample two. As in sample one, the ex post valuations of downloaded albums are nearly a third (27 percent) less than the valuations for purchased albums. The ex ante valuations of downloaded albums, relative to purchased albums, however, are even lower than the



ex post valuations (46 percent vs 27). At the time of obtaining the music, then, persons expect the albums they download to be much less valuable to them than purchased music. This suggests that they would not have been willing to pay much for this music, and that they would not likely have purchased it. But the coefficient estimate describes only the average valuation difference between downloaded and purchased CDs. A more thorough analysis of downloading requires information about the full structure of ex ante valuation (for demand) and ex post valuation (for welfare).

### c. Ex Ante and Ex Post Valuation of Music

This section provides some descriptive characterization of the ex ante and ex post valuation data. For each album, we asked for an ex ante valuation, a price (if they purchased the album), an ex post valuation, as well as indicators for why the current valuation deviated (from the initial valuation). If the current valuation exceeded the initial valuation, we allowed respondents to indicate a) that they were pleasantly surprised, b) that the music grew on them, or c) both. If the current valuation equaled the initial, we allowed respondents to indicate that a) they were familiar with the music before purchase, b) they guessed right, or c) both. Finally, if the current valuation fell short of the initial valuation, we allowed respondents to indicate that a) they were disappointed from the start, b) they grew tired of the music, or c) both.

The data are described in table 11. The average ex ante valuation is \$13.34, while the average ex post valuation is \$11.94 for the 1209 albums in sample two, indicating depreciation. Respondents report higher ex post than ex ante valuations because of pleasant surprise at 16.4 percent of albums and that 16.5 percent of albums grew on them. Respondents report equal ex post and ex ante valuations: 21.8 percent of albums were familiar before purchase, and respondents guessed accurately on the value of 11.5 percent of albums. Finally, respondents report lower ex

post than ex ante valuations for many albums. 9.5 percent of albums are disappointing from the start, while respondents grew tired of 31.7 percent of albums.

These data also allow us to explore the distributions of ex ante and ex post valuations and therefore what the determinants of purchase decisions and welfare look like. The top panel of figure 3 depicts the distribution of ex ante valuations, ordered from highest to lowest ex ante valuations. The curve is, in effect, a demand curve, since points above any price indicate a willingness for an individual to purchase an album. It is worth noting that their demand curve does not appear to belong to any familiar parametric family. If ex ante valuations, which are guesses about how much the person will like an album, are not accurate, then the area under the demand curve does not represent welfare.

The second panel of the figure shows the distribution of ex post valuations, ordered by ex post valuations. This curve is not relevant to purchase decisions, since the valuations are not known at the time of purchase. These valuations are relevant to welfare, however. The curves in the two panels look very similar, but one needs to bear in mind that they are ranked by different variables.

Figure 4 illustrates the distinction between ex ante and ex post valuation more directly. The top panel reproduces the ex ante curve (the line) with ex post valuations overlaid as open circles. It is clear that many purchases with high ex ante value have low ex post value, and vice versa. The bottom panel shows the same information in a different way, with a plot of ex post valuation on ex ante valuation and a 45 degree line. This diagram too shows that, while ex ante and ex post valuations are correlated, the correlation, at 0.63, is far from perfect. Moreover, the preponderance of dots below the 45 degree line indicates depreciation overall.

Column (1) of table 11 reports a regression of ex post valuation on ex ante valuation and dummies that reflect, for example, whether an individual has grown tired of an album. The coefficient on “grew tired” is negative and significant.

The true ex post valuation of an album is the sum over the flow utilities experienced, appropriately discounted. If the flow utility declines over time, then the current valuation of a three year old album will understate its original ex post value. To address this we examine average depreciation directly, via a regression of (ex post valuation – ex ante valuation) on dummies for year of release.<sup>7</sup> Albums released in 1999 are on average \$3 less valuable than albums released in 2003, and the earlier the release, the greater the gap, which we interpret as depreciation. Consequently, we adjust ex post valuation by adding back the average depreciation depicted in column (2) of table 11. By construction this equalizes average ex ante and ex post valuation.

Table 12 presents the average depreciation<sup>8</sup>, along with the averages of the indicators for pleasant surprise, etc, for albums by artists appearing 15 or more times in sample two. The artists are sorted from least to most depreciated. For example, respondents adjusted ex post valuation of the Red Hot Chili Peppers’ record exceeds their ex ante valuation by an average of 31 percent. Nearly a third report being pleasantly surprised by their album, and almost half report that their album grew on them. Albums by the Beatles, Norah Jones, and U2 also appreciated substantially. At the other end of the spectrum, albums by various artists<sup>9</sup> depreciated by an average of 92 percent. Nearly half of respondents with albums by various artists report growing tired of them. Similarly, over 80 percent of Britney Spears album owners report having grown tired of their albums.

---

<sup>7</sup> We don’t know the year the individual obtained the album, but we assume earlier acquisition of albums released sooner.

<sup>8</sup> We calculate depreciation as  $\ln(\text{ex post valuation}/\text{ex ante valuation})$ .

<sup>9</sup> These are primarily albums in the “Now, That’s What I Call Music!” series.

d. Simulating the Effects of Downloading on Sales and Welfare

We now turn to measuring the welfare effects of downloading directly for a subsample.

Given that we have data on ex ante and ex post valuations for respondents in the second Penn sample, we can calculate their expenditure, consumer surplus, and the number of albums purchased and downloaded (in the downloading regime) straightforwardly according to the scheme in table 1.

Assuming that  $p_i$  is \$15, all of the information required to simulate expenditure, consumer surplus, and deadweight loss is available in table 13.<sup>10</sup> Of the 1209 albums in sample 2 obtained under the downloading regime, 617 albums were purchased, while 592 albums were downloaded by the 92 persons in the ex post-ex ante valuation sub-sample. The buyers paid \$15 per album, generating revenue of \$9255. Buyers experienced ex post valuations of \$14.94 per album, giving rise to buyer consumer surplus of -\$37.<sup>11</sup> The 154 downloaded albums that would have been purchased are valued at \$17.91 each, generating \$2738 in surplus. In conjunction with the 438 albums that would not have been purchased and which are ultimately valued at \$9.48 each, the total downloader surplus is \$6910, and the overall consumer surplus under downloading is \$6873.

If downloading were not feasible, the 617 purchased albums would still be purchased, as would 154 downloaded albums with ex ante valuations of \$15 or more, generating revenue of \$11,565 from this sample. Sales and revenue – since prices are assumed constant – are 20 percent lower with downloading than without. Consumer surplus would include both the -\$37 from those

---

<sup>10</sup> We asked prices paid for albums in survey 2. The mean price paid was \$15.30. The simulation results do not change in important ways if we use different prices (for example \$13, 14, or 16).

<sup>11</sup> Note that negative consumer surplus is ruled out by construction with ex ante valuations but is possible with ex post valuation realizations below prices paid.

albums which were purchased along with \$448 from the 154 albums which would have been purchased generating \$17.91 in ex post valuation on average, for a total of -\$411 in consumer surplus. Finally, the sum of the ex post valuations of the albums which would not have been purchased (with ex ante valuations below \$15) generate deadweight losses of \$4152.

In per capita terms, consumers spend \$126 without downloading and \$101 with downloading. Downloading increases consumer welfare by \$70 per capita for sample individuals. Nearly two thirds (64 percent) of this derives from the \$45 per capita reduction in deadweight loss. The remaining 36 percent comes from the \$25 per capita reduction in spending.

If we compute consumer surplus and deadweight loss using ex ante rather than ex post music valuations (see the last part of table 13), the results are similar. Downloading raises consumer welfare for sample individuals by \$64 per capita, and 60 percent of this results from the \$39 per capita reduction in deadweight loss.

#### e. Other Issues

We administered a small-scale follow-up survey to address two questions left unanswered by the previous analysis, the relationship between song and album downloads and the tendency for people to purchase albums following downloading.

Our surveys asked respondents to indicate whether they had obtained “albums” via sharing. As table 5 indicates, Penn students report downloading roughly as many albums as they purchase (about 7.2). Yet, music downloading is done song by song, raising a question of what it means to “download an album.”<sup>12</sup> To address this we administered a follow up survey asking 25 Penn students asking how many songs they had obtained via sharing from each of the albums in the hit

---

<sup>12</sup> When respondents to the earlier asked how to fill out the survey, they were told to include an album if they had obtained they parts they though were valuable to them.

sample. For these students the mean number of purchased albums is 6.96, just slightly below the full Penn sample average. The average number of albums from which they had obtained at least one song via sharing was 19.7, or more than double the number of “albums” they reported obtaining via sharing. On the other hand, they report downloading an average of only 4.0 *entire* albums. It appears, then, that including all albums from which they download 6 or more songs, they obtained 4.1 albums via sharing. Constructing successively more inclusive sharing measures, we get the following: including 5 or more songs yields 4.6; 4 or more yields 6.3; 3 or more yields 8.8; and 2 or more yields 14.3. It therefore appears that respondents reported obtaining an album if they have downloaded 3 or 4 (or more) of its songs.

The follow-up survey also allows us to examine whether people download, then buy. Of the 476 albums – or album fragments – obtained via sharing, respondents subsequently purchased 21, or 4.4 percent.

## V. Conclusion

We argue that successfully measuring the possible sales-displacing effect of unpaid music downloading requires data on the quantities of purchases and downloads made by individuals, leading us to conduct original surveys. Using a variety of empirical approaches, we document that downloading displaces sales among a convenience sample of college students. The estimate we consider most conservative indicate that an additional download reduces sales by between 0.1 and 0.2 units. As a result, for the individuals in our sample, downloading reduced expenditure by about 10 percent but possibly much more. Supporting incomplete sales displacement is our finding that downloaded music is valued much less than purchased music.

While downloading reduces expenditure (on hit albums, 1999-2003) by \$25 per capita in the sub-sample for which we perform a direct welfare analysis of downloading, it raises sample consumers' welfare associated with these albums by \$70 per capita. Some of the benefit to consumers are transfers from sellers, but most of the benefit (\$45 per capita) comes from reductions in deadweight loss.

Two facts bear emphasis again. First, our sample is not representative, so our results should not be generalized. Second, our evaluation of welfare takes supply as given. It is entirely possible that downloading has important effects on the quantity and types of music recorded and marketed in the first place. This is an important area for further research.

## References

Bakos, Yannis, Erik Brynjolfsson, and Douglas Lichtman. "Shared Information Goods." *Journal of Law & Economics* XLII (April 1999): 117-155.

Besen, Stanley M. "Private Copying, Reproduction Costs, and the Supply of Intellectual Property," *Information Economics and Policy*, 2 (1986): 5-22.

Conner, Kathleen Reavis and Richard Rumelt. "Software Piracy: An Analysis of Protection Strategies." *Management Science*, vol 37, no. 2 (Feb 1991): 125-139.

Diamond, Peter A. and Jerry A. Hausman. "Contingent Valuation: Is Some Number Better than No Number?" *Journal of Economic Perspectives*, vol. 8, no. 4, Fall 1994, pp. 45-64

Dixit, Avinash K; Stiglitz, Joseph E. "Monopolistic Competition and Optimum Product Diversity." *American Economic Review*. Vol. 67 (3). p 297-308. June 1977.

Fader, Peter S. Expert Report, *A&M Records, Inc. v. Napster, Inc.* 114 F Supp. 2d 896, 2000.

Graves, Tom. "Standard & Poor's Industry Surveys: Movies & Home Entertainment." Vol. 172, no. 7, section 1. February 12, 2004.

Hui, Kai-Leung and Ivan Png. "Piracy and the Legitimate Demand for Recorded Music." *Contributions to Economic Analysis*. Vol. 2, no. 1, article 11 (2003).

Liebowitz, Stanley J. "Copying and Indirect Appropriability." *Journal of Political Economy*, vol. 93, No. 5 (Oct 1985): 945-957

Liebowitz, Stan "Will MP3 Downloads Annihilate the Record Industry? The Evidence So Far." Unpublished Manuscript. University of Texas at Dallas. June 2003.

Oberholzer-Gee, Felix and Koleman Strumpf. "The Effect of File Sharing on Record Sales: An Empirical Analysis." Mimeo. Harvard Business School. January 2004.

Shapiro, Carl and Hal Varian. *Information Rules*. Harvard Business School Press. Cambridge, MA. 1999.

Spence, Michael. "Product Selection, Fixed Costs, and Monopolistic Competition." *The Review of Economic Studies*. Vol. 43 (2). p 217-35. June 1976.

Spence, Michael. "Product Differentiation and Welfare." *American Economic Review*. Vol. 66 (2). p 407-14. May 1976.

Takeyama, Lisa. "The Intertemporal Consequences of Unauthorized Reproduction of Intellectual Property." *Journal of Law & Economics*, vol. XL, (October 1997): 511-522.



Varian, Hal. "Buying, Sharing, and Renting Information Goods." *Journal of Industrial Economics*. Vol. XLVIII, no. 4, Dec 2000, pp. 473-488

Zentner, Alejandro. "Measuring the Effect of Online Piracy on Music Sales." Unpublished Manuscript. University of Chicago. December 2003.

Table 1: Summary of Behavior and Welfare in Downloading and Non-Downloading Regimes

1. DOWNLOADING REGIME (Illegal downloading is feasible)

Ex ante valuations relative to price:	Willing to download ( $d_i = 1$ )	Not willing to download ( $d_i = 0$ )
$0 < v_i^a < p_i$	<ul style="list-style-type: none"> <li>• <math>v_i^p = \text{CS}</math></li> <li>• <math>\text{rev}=0</math></li> <li>• downloading occurs</li> </ul>	<ul style="list-style-type: none"> <li>• <math>v_i^p = \text{DWL}</math></li> <li>• <math>\text{rev}=0</math></li> <li>• album not obtained</li> <li>• not observed in our data</li> </ul>
$v_i^a \geq p_i$	<ul style="list-style-type: none"> <li>• <math>v_i^p = \text{CS}</math></li> <li>• <math>\text{rev}=0</math></li> <li>• downloading occurs</li> </ul>	<ul style="list-style-type: none"> <li>• <math>v_i^p - p_i = \text{CS}</math></li> <li>• <math>p_i = \text{revenue}</math></li> <li>• purchase occurs</li> </ul>

2. NON-DOWNLOADING REGIME (Illegal downloading is not feasible)

Ex ante valuations relative to price:	Would have been willing to download ( $d_i = 1$ )	Would not have been willing to download ( $d_i = 0$ )
$0 < v_i^a < p_i$	<ul style="list-style-type: none"> <li>• <math>v_i^p = \text{DWL}</math></li> <li>• <math>\text{rev}=0</math></li> <li>• album not obtained</li> </ul>	<ul style="list-style-type: none"> <li>• <math>v_i^p = \text{DWL}</math></li> <li>• <math>\text{rev}=0</math></li> <li>• album not obtained</li> <li>• not observed in our data</li> </ul>
$v_i^a \geq p_i$	<ul style="list-style-type: none"> <li>• <math>v_i^p - p_i = \text{CS}</math></li> <li>• <math>p_i = \text{revenue}</math></li> <li>• purchase occurs</li> </ul>	<ul style="list-style-type: none"> <li>• <math>v_i^p - p_i = \text{CS}</math></li> <li>• <math>p_i = \text{revenue}</math></li> <li>• purchase occurs</li> </ul>

Note: Assume that zero marginal cost distribution is available when illegal downloading is not feasible.

Table 2: Respondent Characteristics

	overall	CCNY	Chicago	Hunter	Penn Survey 1	Penn Survey 2
Male	60% of 414	65% of 66	29.7% of 64	45.4% of 33	67.7% of 251	58% of 98
Black	8.9	31.8	10.9	6.3	4.0	3.0
Hispanic	9.8	33.3	3.1	25.0	3.2	1.0
Asian	31.8	21.2	26.6	12.5	38.5	43.4
White	48.2	13.6	59.4	53.1	53.8	52.5
N	409	66	64	32	247	99
Age	21.9/400	24.6	27.4	20.5	20.0	20.2
<\$25	14.3%	23	41.4	12.9	5.2	2.2
\$25-50k	19.8%	41.5	25.9	38.7	19.8	8.8
\$50-100	27.6%	29.2	17.2	29.3	27.6	36.3
\$100-250	21.9%	4.6	12.1	12.9	30.4	34.1
>\$250	16.4%	1.5	3.5	6.5	24.8	18.7
N	384	65	58	31	230	91
Fandom 1(low)	6.3	6.2	6.3	3.0	6.8	11.1
2	9.2	9.2	10.9	6.1	9.2	10.1
3	38.4	47.7	43.8	30.3	35.6	33.3
4	29.6	18.5	12.3	36.4	33.2	29.3
5	16.5	18.5	15.6	24.2	15.2	16.2
Dialup '99	70.4%	85.2%	56.5%	Na	73.2%	76.0
'00	63.1%	82.1%	60.3%	Na	60.9%	59.8
'01	49.8%	77.8%	58.7%	Na	42.0%	29.9
'02	31.5%	63.0%	47.6%	Na	20.6%	20.6
'03	16.2%	53.6%	32.3%	Na	4.1%	4.1
Present	15.2%	57.1%	30.6%	Na	2.4%	3.2
Hi-speed '99	20.6%	0.0%	29.0%	Na	20.8%	17.7
'00	31.9%	7.1%	28.6%	Na	37.3%	38.1
'01	47.1%	18.5%	31.7%	Na	57.4%	69.1
'02	66.5%	33.3%	46.0%	Na	79.4%	79.4
'03	83.8%	46.4%	67.7%	Na	95.9%	95.9
Present	86.0%	42.9%	69.4%	Na	97.6%	96.8
N	260	28	63		170	97
Collection	103.3	157.9	147.5	96.4	79.0	83.6

Table 3: Average Value of Hits and Current Purchases

	hit			current		
	Mean	median	N	Mean	median	N
SURVEY 1 (buy)						
Ex post						
bought	12.70	12	2726	15.25	15	695
downloaded	8.81	8.5	1512	11.15	10	337
SURVEY 1 (sell)						
Ex post						
bought	56,344	100	702	1,831	100	173
downloaded	2,534	30	264	710	50	64
SURVEY 2						
Ex ante						
bought	15.91	15	617			
downloaded	10.66	10	592			
Ex post						
bought	13.39	13	617			
downloaded	10.47	10	592			

Table 4: Top 40 Sample Albums

artist/album	sample freq	sample rank	sales (mil)	sales rank
50 CENT / GET RICH OR DIE TRYIN'	112	1	5	34
EMINEM / THE EMINEM SHOW	103	2	8	10
COLDPLAY / A RUSH OF BLOOD TO THE HEAD	100	3	2	102
EMINEM / THE MARSHALL MATHERS LP	88	4	8	11
EMINEM / THE SLIM SHADY LP	87	5	4	45
RED HOT CHILI PEPPERS / CALIFORNICATION	77	6	5	35
2 PAC / GREATEST HITS	74	7	9	8
SANTANA / SUPERNATURAL	69	8	14	1
U2 / ALL THAT YOU CAN'T LEAVE BEHIND	68	9	3	59
BEATLES, THE / 1	66	10	8	12
MAYER, JOHN / ROOM FOR SQUARES	60	11	3	60
JONES, NORAH / COME AWAY WITH ME	57	12	7	19
DMX / FLESH OF MY FLESH, BLOOD OF MY BLO	56	13	3	61
MATTHEWS, DAVE BAND / EVERYDAY	56	14	3	62
'N SYNC / NO STRINGS ATTACHED	53	15	11	4
BACKSTREET BOYS / MILLENNIUM	52	16	13	2
JAY-Z / THE BLUEPRINT	52	17	2	103
NELLY / COUNTRY GRAMMAR	52	18	8	13
CAREY, MARIAH / #1'S	50	19	5	36
DMX / ...AND THEN THERE WAS X	49	20	5	37
DR. DRE / DR. DRE 2001	49	21	6	23
LINKIN PARK / HYBRID THEORY	49	22	8	14
OUTKAST / STANKONIA	48	23	3	63
BLINK 182 / ENEMA OF THE STATE	47	24	5	38
MATCHBOX TWENTY / MAD SEASON	47	25	4	46
CREED / HUMAN CLAY	46	26	10	6
DIDO / NO ANGEL	46	27	4	47
DOORS, THE / GREATEST HITS (1996)	45	28	2	104
MADONNA / MUSIC	42	29	2	105
JAY-Z / THE BLUEPRINT 2: THE GIFT AND TH	41	30	3	64
DESTINY'S CHILD / SURVIVOR	40	31	4	48
JAY-Z / VOL. 3...LIFE AND TIMES OF S. CA	40	32	3	65
SPEARS, BRITNEY / ...BABY ONE MORE TIME	40	33	13	3
'N SYNC / CELEBRITY	39	34	5	39
ENYA / A DAY WITHOUT RAIN	39	35	6	24
NOTORIOUS B.I.G. / BORN AGAIN	39	36	2	106
SPEARS, BRITNEY / OOPS!...I DID IT AGAIN	39	37	9	9
TIMBERLAKE, JUSTIN / JUSTIFIED	39	38	3	66
2 PAC / UNTIL THE END OF TIME	37	39	3	67
MOBY / PLAY	37	40	2	107

Table 5: Purchases and Downloads

	Hit	hit	hit	current	current	current
	purchase	download	N	purchase	download	N
All	8.24	5.36	365	2.61	1.24	349
Asian	6.74	7.98	108	1.98	1.16	101
Black	9.53	2.34	32	3.34	1.31	32
Hispanic	11.62	4.83	37	2.85	1.03	33
White	8.31	4.55	184	2.75	1.35	179
CCNY	10.96	3.24	54	2.90	0.63	52
Chicago	8.21	2.25	56	2.82	0.54	57
Hunter	11.13	1.72	32	3.96	0.56	27
Penn	7.17	7.17	365	2.31	1.67	213

Table 6: Hit Purchases and Downloads, by Year of Albums' First Certification

		total	Total	CCNY	CCNY	Chicago	Chicago	Hunter	Hunter	Penn	Penn
	% of hits downloaded	# purchased	# dwnloaded	# purchased	# dwnloaded	# purchased	# dwnloaded	# purchased	# dwnloaded	# purchased	# dwnloaded
1999	32.9%	2.39	1.17	2.90	0.74	1.78	0.41	2.88	0.38	2.35	1.57
2000	38.9%	2.43	1.55	3.32	0.92	2.33	0.69	3.34	0.53	2.12	2.07
2001	41.7%	1.76	1.26	2.86	0.82	1.91	0.44	2.41	0.47	1.38	1.68
2002	45.8%	1.28	1.08	1.60	0.52	1.61	0.67	1.43	0.19	1.09	1.45
2003	46.4%	0.52	0.45	1.10	0.42	0.31	0.13	1.06	0.16	0.36	0.58

Table 7: OLS and Tobit Sales Displacement Estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Hit Purchases	Hit Purchases	Hit Purchases	Current Purchases	Current Purchases	Current Purchases	2003 Hit Purchases
Current Downloads				-0.1945	-0.1722	-0.2579	
				(0.0718)**	(0.0728)*	(0.1001)*	
Hit Downloads	-0.0266	-0.0085	-0.0310				-0.0910
	(0.0521)	(0.0527)	(0.0598)				(0.0556)
male	-3.3089	-3.5692	-4.4627	-0.6344	-0.6188	-1.0338	
	(0.9708)**	(1.0173)**	(1.1373)**	(0.3167)*	(0.3306)	(0.4419)*	
Fan = 2	4.0584	3.7992	8.0407	0.1022	0.0572	1.1415	0.3705
	(2.4907)	(2.4789)	(2.9994)**	(0.8161)	(0.8161)	(1.2328)	(0.2494)
Fan = 3	5.6050	5.3117	9.2048	0.9109	0.8855	2.4968	0.5430
	(2.1314)**	(2.1219)*	(2.6438)**	(0.6982)	(0.6979)	(1.0789)*	(0.2128)*
Fan = 4	6.7212	6.4736	10.6962	1.9665	1.9228	3.9187	0.6201
	(2.1738)**	(2.1652)**	(2.6787)**	(0.7144)**	(0.7141)**	(1.0955)**	(0.2173)**
Fan = 5	9.9391	9.5997	14.1591	2.6707	2.5686	4.8472	0.6221
	(2.3228)**	(2.3144)**	(2.8243)**	(0.7518)**	(0.7527)**	(1.1391)**	(0.2327)**
\$25-50k	2.2034	1.6273	2.0517	0.5603	0.4071	0.4576	0.0848
	(1.5629)	(1.5824)	(1.7785)	(0.4988)	(0.5089)	(0.6768)	(0.1564)
\$50-100	1.0381	1.1821	1.5687	0.2371	0.2931	0.1194	-0.0025
	(1.3790)	(1.4304)	(1.6202)	(0.4659)	(0.4823)	(0.6504)	(0.1373)
\$100-250	0.7783	1.3815	2.2364	0.3413	0.5042	0.5354	0.0577
	(1.4597)	(1.5420)	(1.7435)	(0.4827)	(0.5088)	(0.6845)	(0.1450)
>\$250	2.6972	3.4626	4.3480	0.4111	0.6026	0.5559	0.0485
	(1.6185)	(1.7117)*	(1.9276)*	(0.5340)	(0.5647)	(0.7550)	(0.1606)
black	1.7491	0.3582	0.6459	0.3397	0.2338	0.4188	0.7495
	(1.7698)	(1.8678)	(2.0700)	(0.5682)	(0.5923)	(0.7799)	(0.1768)**
hisp	1.7620	-0.0623	0.0169	-0.2467	-0.5010	-0.2740	0.2155
	(1.6841)	(1.8210)	(2.0131)	(0.5553)	(0.6063)	(0.7901)	(0.1689)
asian	-0.8649	-0.9258	-1.6524	-0.7525	-0.6987	-1.0400	0.2989
	(1.1388)	(1.1535)	(1.3002)	(0.3611)*	(0.3667)	(0.4964)*	(0.1126)**
Chicago		-3.5870	-3.9167		-0.3204	-0.5408	
		(1.9267)	(2.1508)		(0.6088)	(0.8167)	
Hunter		-1.4116	-1.3880		0.4883	0.6874	
		(2.0773)	(2.3001)		(0.7008)	(0.9171)	
Penn		-4.1112	-4.5066		-0.6275	-0.5965	
		(1.6526)*	(1.8442)*		(0.5351)	(0.7167)	
	OLS	OLS	tobit	OLS	OLS	tobit	OLS
Constant	2.8597	6.4934	2.1786	1.7047	2.0768	0.0260	-0.1857
	(2.2992)	(2.8164)*	(3.3257)	(0.7482)*	(0.9027)*	(1.3107)	(0.2214)
Observations	364	364	364	347	347	347	364
R-squared	0.12	0.14		0.14	0.15		

Standard errors in parentheses. \* significant at 5%; \*\* significant at 1%.



Table 8: IV Estimates of Hit and Current Sales Displacement

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample	Hit	Hit	Hit	Hit	Current	Current	Current	Current
Depend. Var	Download	Download	Purchase	Purchase	Download	Download	Purchase	Purchase
Instrument	School	Individ.	School	Individ.	School	Individ.	School	Individ.
Current Downloads							-0.8804	-1.4676
							(0.4570)	(0.8502)
Hit Downloads			-0.5719	-0.2563				
			(0.3334)	(0.4415)				
Broadband Access		3.9311				1.0707		
		(2.0540)				(0.5222)*		
male	0.1042	0.4737	-2.9788	-3.0819	0.5737	0.9173	-0.2050	0.4658
	(1.0355)	(1.3606)	(1.1299)**	(1.1891)*	(0.2475)*	(0.3465)**	(0.4551)	(1.0341)
Fan =2	3.3970	3.9291	5.8858	6.4379	0.6286	0.7874	0.4870	0.7147
	(2.5165)	(3.3562)	(3.0582)	(3.1422)*	(0.6149)	(0.8373)	(0.9552)	(1.5463)
Fan =3	6.0806	7.5175	8.8921	6.8162	0.7307	0.9770	1.3940	2.3849
	(2.1350)**	(2.7380)**	(3.1421)**	(3.9991)	(0.5251)	(0.7038)	(0.8494)	(1.4629)
Fan =4	5.8373	6.7521	9.8673	8.7015	0.8491	1.0916	2.5449	3.2605
	(2.1814)**	(2.8182)*	(3.1280)**	(3.7107)*	(0.5368)	(0.7232)	(0.8912)**	(1.5214)*
Fan =5	4.8969	5.3952	12.4919	9.9086	0.8452	1.3516	3.1924	4.8208
	(2.3410)*	(3.0554)	(3.0726)**	(3.4224)**	(0.5661)	(0.7707)	(0.9150)**	(1.7036)**
\$25-50k	0.6323	1.1859	2.3412	3.5615	0.1678	-0.1327	0.6081	-0.0888
	(1.6102)	(2.2458)	(1.7926)	(1.9947)	(0.3839)	(0.5638)	(0.5640)	(0.9854)
\$50-100	1.3031	1.8318	2.2499	1.5951	0.5565	0.4007	0.7611	1.0134
	(1.4543)	(1.9315)	(1.7400)	(1.9552)	(0.3627)	(0.5105)	(0.6282)	(1.0181)
\$100-250	1.6783	0.9198	2.5392	0.1394	0.6131	0.4236	1.0042	0.9678
	(1.5669)	(2.0122)	(1.9796)	(1.8480)	(0.3825)	(0.5225)	(0.6970)	(1.0653)
>\$250	0.9148	0.4349	4.1691	3.4247	0.6004	0.2875	1.1005	0.6888
	(1.7415)	(2.2544)	(2.0549)*	(1.9585)	(0.4248)	(0.5968)	(0.7535)	(1.1256)
black	-0.4683	0.1327	1.0382	0.8998	0.6541	0.8319	0.6239	0.3569
	(1.9010)	(2.5156)	(2.0723)	(2.1152)	(0.4455)	(0.6096)	(0.6679)	(1.2014)
hispanic	2.2792	4.3455	2.3844	4.9284	0.4618	0.9973	-0.1939	1.1861
	(1.8495)	(2.6193)	(1.9655)	(2.5741)	(0.4568)	(0.6604)	(0.6278)	(1.2501)
asian	3.4776	5.8701	1.3554	0.7168	-0.0192	0.1463	-0.7229	-0.9885
	(1.1592)**	(1.5129)**	(1.8671)	(2.8829)	(0.2767)	(0.3935)	(0.4081)	(0.6960)
Chicago	0.0573				0.4089			
	(1.9611)				(0.4588)			
Hunter	-1.3440				0.0841			
	(2.1131)				(0.5288)			
Penn	3.6365				1.0906			
	(1.6707)*				(0.3994)**			
Constant	-4.1479	-6.7023	1.2116	1.7005	-1.0581	-1.4387	1.3955	1.1489
	(2.8580)	(3.3334)*	(2.8147)	(2.9569)	(0.6787)	(0.8401)	(0.8686)	(1.3918)
Observations	364	231	364	231	347	219	347	219

Notes: Standard errors in parentheses. \* significant at 5%; \*\* significant at 1%

Table 9: Longitudinal Sales Displacement Estimates  
(hit sample)

	(1)	(2)	(3)
	Purch's from yr.	Purch's from yr. all	Purch's from yr. >5
Dwnld's from yr.	-0.0447 (0.0458)	-0.0789 (0.0334)*	-0.1565 (0.0656)*
d2000	0.0527 (0.1138)	0.0656 (0.1137)	0.2061 (0.2113)
d2001	-0.6141 (0.0995)**	-0.6110 (0.1131)**	-0.9055 (0.2096)**
d2002	-1.1111 (0.1185)**	-1.1141 (0.1131)**	-2.0608 (0.2096)**
d2003	-1.8747 (0.1291)**	-1.8987 (0.1154)**	-3.5251 (0.2135)**
Fan =2	0.9301 (0.2558)**		
Fan =3	1.2497 (0.2035)**		
Fan =4	1.4370 (0.2221)**		
Fan =5	2.0663 (0.3374)**		
\$25-50k	0.3723 (0.2932)		
\$50-100	0.1326 (0.2475)		
\$100-250	0.0702 (0.2776)		
>\$250	0.3923 (0.2710)		
black	0.4463 (0.3307)		
hisp	0.4105 (0.3766)		
asian	-0.0662 (0.2273)		
Constant	0.8397 (0.2218)**	2.4469 (0.0885)**	4.4216 (0.1625)**
Observations	1820	1820	820
R-squared	0.14	0.22	0.39
method	OLS	FE	FE
Number of individuals		364	164

Notes: Robust standard errors in parentheses. \* significant at 5%; \*\* significant at 1%.

Table 10: Relative Valuation of Downloaded CDs

a. Survey One – Ex Post Valuation (Hit and Current Samples)

	(1)	(2)	(3)	(4)
	Log value hit	Log value current	Log value hit	Log value current
Downloaded	-0.330 (0.072)**	-0.390 (0.089)**	-0.177 (0.023)**	-0.238 (0.041)**
Constant	2.382 (0.037)**	2.605 (0.038)**	2.331 (0.011)**	2.557 (0.018)**
Observations	4017	1011	4017	1011
R-squared	0.05	0.08	0.02	0.04
	OLS	OLS	FE	FE
Number of individuals			284	226

Notes: Robust standard errors in parentheses. \* significant at 5%; \*\* significant at 1%

b. Survey Two – Ex Ante and Ex Post Valuations (Hits Only)

	(1)	(2)	(3)	(4)
	Log Ex Ante Valuation	Log Ex Post Valuation	Log Ex Ante Valuation	Log Ex Post Valuation
Downloaded	-0.4623 (0.0822)**	-0.2654 (0.1043)*	-0.2857 (0.0283)**	-0.0857 (0.0511)
Constant	2.7104 (0.0348)**	2.3879 (0.0765)**	2.6253 (0.0175)**	2.3013 (0.0317)**
Observations	1133	1133	1133	1133
R-squared	0.15	0.02	0.09	0.00
	OLS	OLS	FE	FE
Number of individuals			93	92

Notes: Robust standard errors in parentheses. \* significant at 5%; \*\* significant at 1%

Table 11: Ex Ante and Ex Post Valuations

		(1)	(2)
	mean	Ex Post Valuation	Ex Post – Ex Ante Valuation
Ex Ante Valuation	\$13.34	0.8178 (0.0183)**	
Ex Post Valuation	\$11.94		
ex ante and ex post correlation	0.6389		
Pleasantly surprised	16.4%	4.2537 (0.4657)**	
Grew on me	16.5%	4.7717 (0.4656)**	
Familiar before I bought	21.8%	0.2707  (0.4921)	
Guessed right	11.5%	0.1277 (0.5110)	
Disappointed from start	9.5%	-5.5276  (0.5782)**	
Grew tired of it	31.7%	-5.2951 (0.5028)**	
Released in 99			-3.1906 (0.7635)**
Released in 00			-1.7907 (0.7534)*
Released in 01			-1.7723 (0.7931)*
Released in 02			-0.0587 (0.8038)
Constant		1.6733 (0.5214)**	0.3375 (0.6831)
Observations		1209	1209
R-squared		0.74	0.03

Notes: Standard errors in parentheses. \* significant at 5%; \*\* significant at 1%

Table 12: Album Depreciation for Major Artists

artist	adjusted depreciation	Pleasantly surprised	Grew on me	Familiar before got it	Gessed right	Disappointed from start	Got tired of it	N
RED HOT CHILI PEPPERS	31.3%	30.0%	43.3%	26.7%	10.0%	0.0%	13.3%	30
BEATLES, THE	26.2%	25.9%	22.2%	33.3%	0.0%	0.0%	22.2%	27
JONES, NORAH	17.4%	35.3%	58.8%	11.8%	5.9%	11.8%	11.8%	17
U2	9.8%	18.8%	43.8%	18.8%	6.3%	6.3%	18.8%	16
LINKIN PARK	9.8%	37.0%	14.8%	29.6%	7.4%	7.4%	14.8%	27
DION, CELINE	9.7%	18.8%	0.0%	25.0%	12.5%	12.5%	31.3%	16
COLDPLAY	6.3%	25.0%	27.8%	27.8%	22.2%	0.0%	8.3%	36
2 PAC	1.9%	10.7%	25.0%	21.4%	7.1%	0.0%	42.9%	28
EMINEM	0.7%	22.9%	21.4%	21.4%	8.6%	0.0%	31.4%	70
SOUNDTRACK	0.0%	28.3%	23.9%	19.6%	6.5%	6.5%	23.9%	45
DOORS, THE	-1.2%	12.5%	12.5%	25.0%	12.5%	6.3%	37.5%	16
MATTHEWS, DAVE BAND	-5.2%	7.1%	19.0%	26.2%	14.3%	19.0%	26.2%	42
MOBY	-6.2%	5.9%	5.9%	47.1%	5.9%	17.6%	17.6%	17
MAYER, JOHN	-8.7%	31.3%	25.0%	18.8%	6.3%	6.3%	25.0%	16
NELLY	-17.2%	15.8%	21.1%	15.8%	5.3%	5.3%	31.6%	19
DESTINY'S CHILD	-20.2%	0.0%	6.7%	20.0%	20.0%	6.7%	46.7%	15
50 CENT	-20.5%	17.4%	21.7%	17.4%	8.7%	0.0%	39.1%	23
'N SYNC	-20.7%	4.8%	0.0%	4.8%	4.8%	14.3%	71.4%	21
AGUILERA, CHRISTINA	-21.6%	10.3%	0.0%	10.3%	13.8%	10.3%	55.2%	29
BLINK 182	-23.0%	6.7%	13.3%	6.7%	6.7%	26.7%	46.7%	15
CAREY, MARIAH	-23.5%	0.0%	10.5%	26.3%	5.3%	31.6%	26.3%	19
BACKSTREET BOYS	-24.5%	0.0%	0.0%	17.6%	5.9%	11.8%	64.7%	17
DMX	-24.8%	10.5%	5.3%	21.1%	10.5%	10.5%	52.6%	19
SPEARS, BRITNEY	-28.3%	0.0%	3.4%	6.9%	3.4%	3.4%	82.8%	29
JA RULE	-48.4%	6.7%	0.0%	26.7%	0.0%	26.7%	40.0%	15
JAY-Z	-52.4%	17.4%	0.0%	8.7%	17.4%	26.1%	34.8%	23
VARIOUS	-86.9%	10.0%	0.0%	5.0%	5.0%	10.0%	45.0%	15

618

Note: “depreciation” is  $\log(\text{adjusted ex post value}/\text{ex ante value})$ . Artists appear in this table if their albums appear 15+ times in the sample for survey 2.

Table 13: Revenue and Welfare with and without Downloading

	N	price	ex ante valuation	ex post valuation	ex post valuation (adjusted)	
did buy	617	15	15.91	13.34	14.94	
<u>Downloaded, but:</u>						
would buy if downloading were not feasible	154	15	17.82	16.72	17.91	
would not buy	438	Na	8.15	8.28	9.48	
	No Downloading		Downloading		Change	
	total	per capita	total	per capita	Total	per capita
Quantity sold	771	8.38	617	6.71	-154	-1.67
Quantity Downloaded	0	0	592	6.43	592	
Quantity Consumed	771	8.38	1209	13.41	438	4.76
Revenue	\$11,565	\$126	\$9,255	\$101	-\$2,310	-\$25.1
CS and DWL (ex post valuations)						
Buyer CS	\$411	\$4.47	-\$37	-\$0.40	-\$448	-\$4.87
Downloader CS	\$0	\$0	\$6,910	\$75.1	\$6,910	\$75.1
CS total	\$411	\$4.47	\$6,873	\$74.7	\$6,462	\$70.2
DWL*	\$4,152	\$45.1	\$0	\$0	-\$4152	-\$45.1
CS and DWL (ex ante valuations)						
Buyer CS	\$996	\$10.82	\$561	\$6.10	-\$434	-\$4.72
Downloader CS	\$0	\$0.00	\$6,314	\$68.6	\$6,314	\$68.6
CS total	\$996	\$10.82	\$6,875	\$74.7	\$5,880	\$63.9
DWL	\$3,570	\$38.80	\$0	\$0.00	-\$3,570	-\$38.8

Notes: Based on 92 individuals in survey 2 sample. We assume that the price of albums is \$15. Welfare calculations based on ex post valuation adjusted for depreciation. We assume that a downloaded album would have been purchased if the consumer's ex ante valuation  $\geq$  \$15.

\* DWL in this table excludes foregone welfare from albums that consumers are unwilling to either purchase or download. See table 1 and the surrounding text.

Figure 1: Shipments, 1993-2003

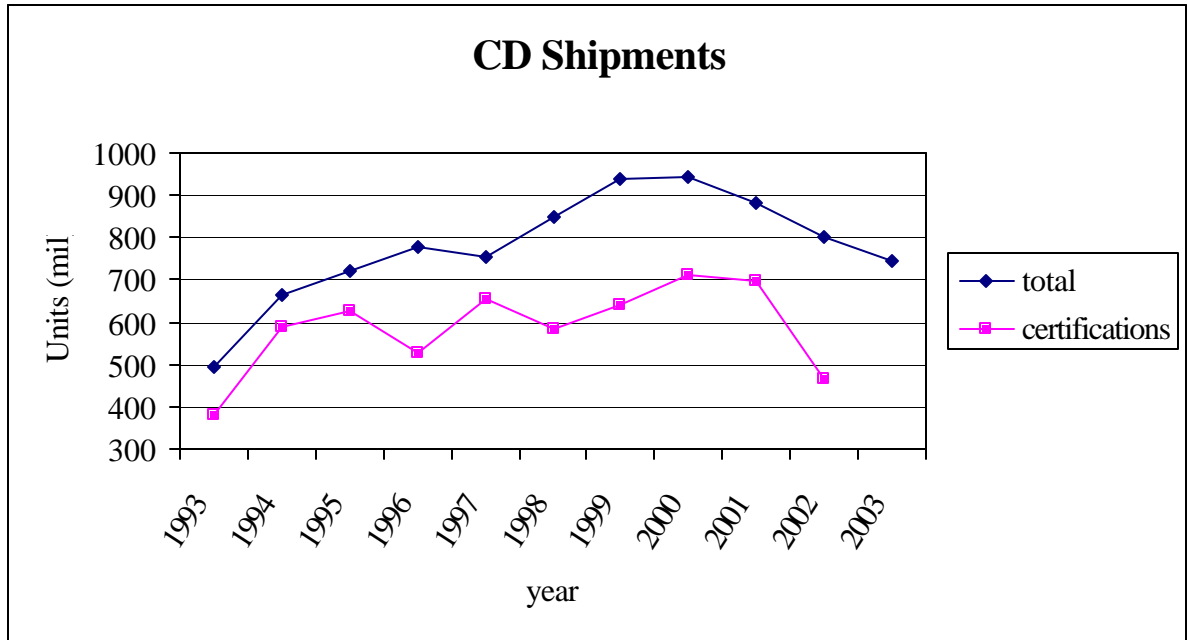
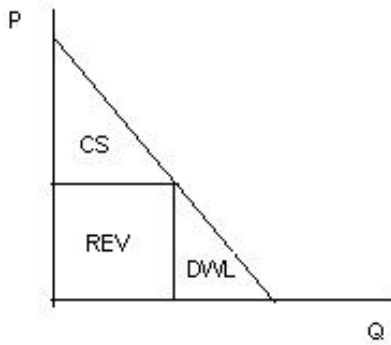
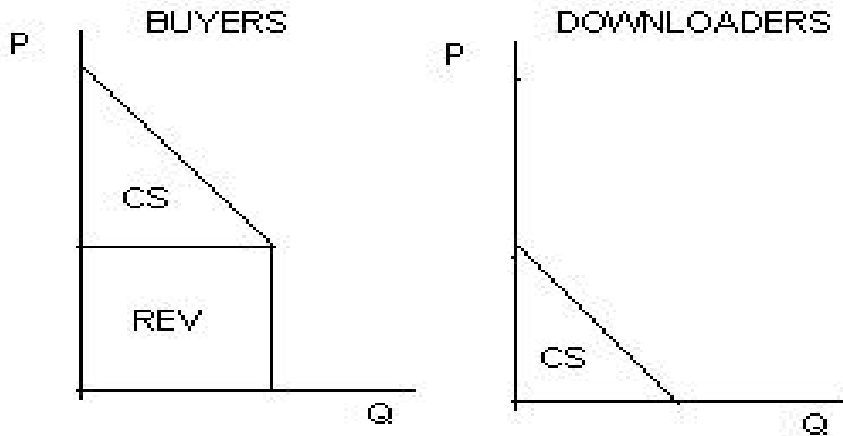


Figure 2: Effects of Downloading on Welfare

1. single-price monopoly - the market before downloading



2. downloaders are low-valuation demanders



3. downloaders are high-valuation demanders

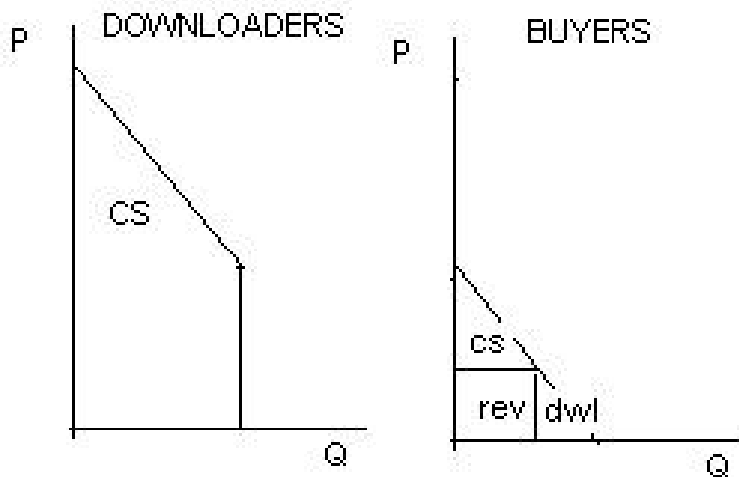




Figure 3: Ex Ante Valuation (Demand) and Ex Post Valuation (Welfare)

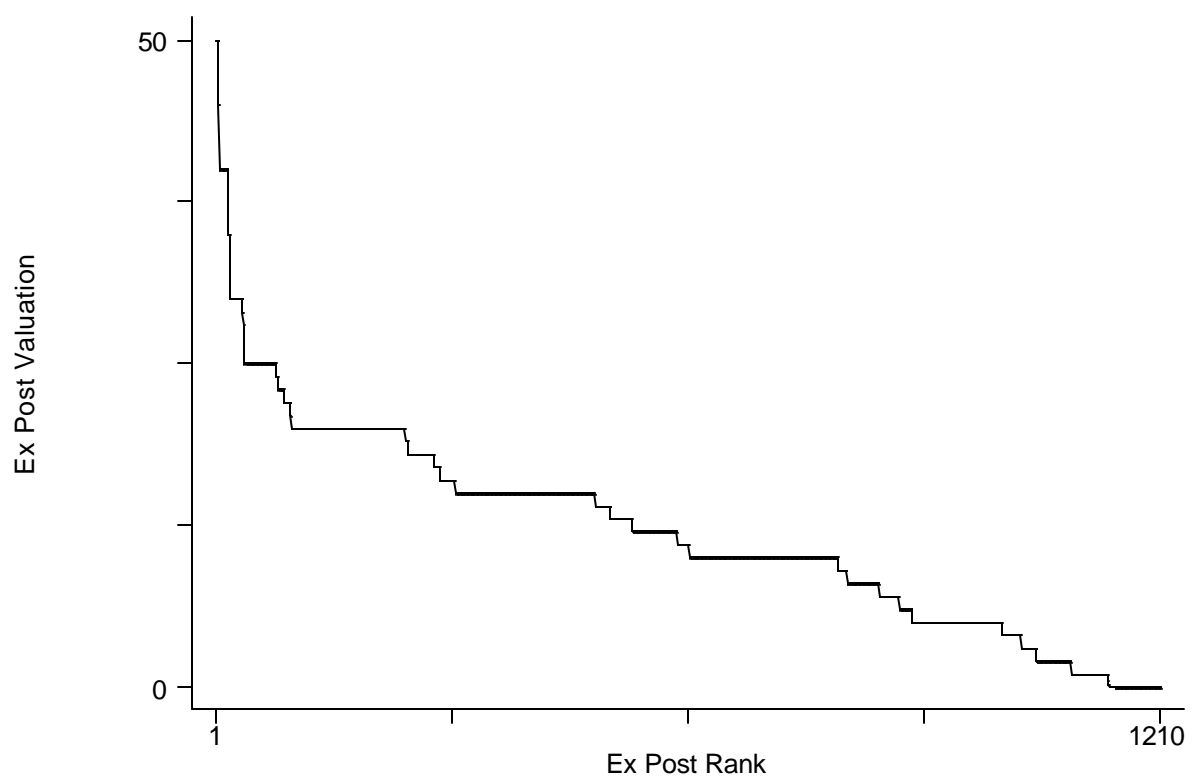
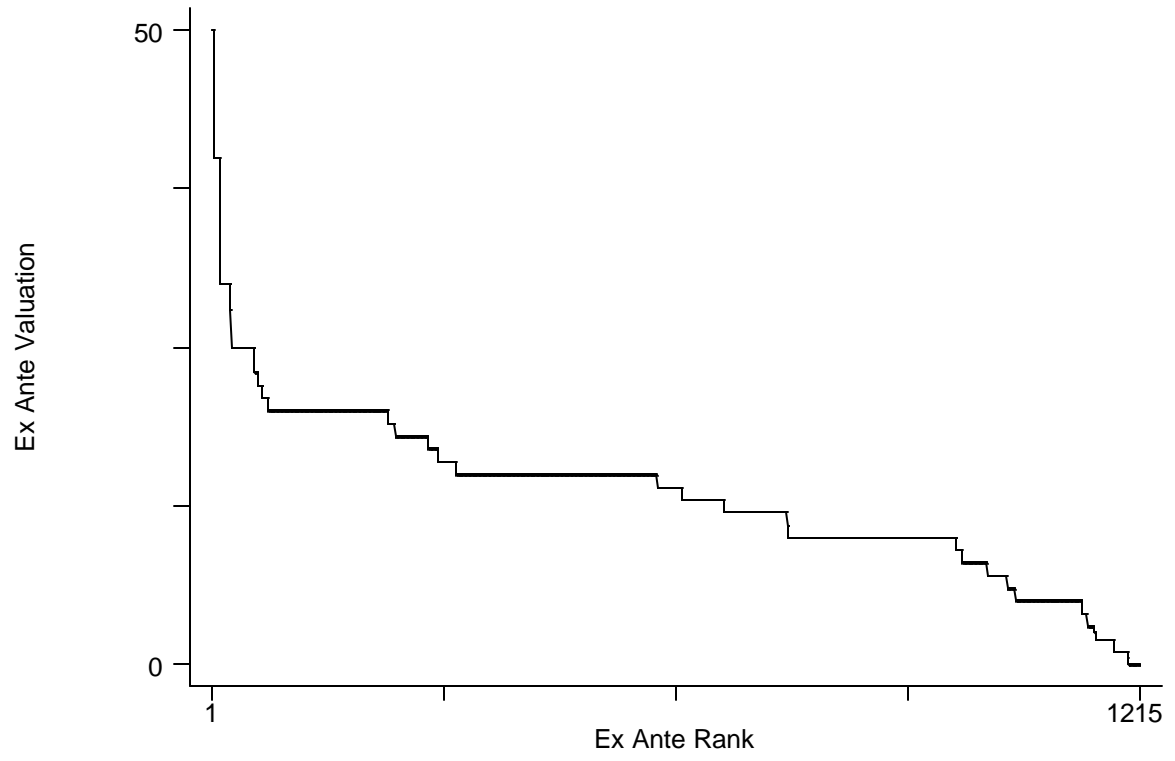


Figure 4: The Demand Curve and Ex Post Valuation

