Developing Superstars: The Effect of Unauthorized Copying on Investment in Musical Talent

David Blackburn*

July 27, 2005

Abstract

While much attention has been focused on the short-run effects that unauthorized copying (file sharing) has had on the sales of recorded music, little attention has been paid to the long-run implications of these effects. In fact, the issues facing the music industry present a trade-off between short-run efficiency and long-run investment incentives similar to that in the patent literature. We focus on investigating how the changes to the distribution of album sales brought on by file sharing, which are detailed in Blackburn (2004), has shifted incentives for record companies to invest in developing artists. Attention is paid both to the effects on investment in new artists, as well as the investments made to develop new works from already established artists. A model is developed which illustrates that while file sharing certainly reduces incentive to release albums from already established artists, there is an increased incentive to develop new artists. The implications of this model are tested with data from Billboard magazine's weekly Hot 200 sales chart over a 14 year period from 1992 to 2005. Though the evidence is somewhat mixed, there is little to suggest that there have been major changes on the supply side of the recorded music industry since 1999.

*Harvard University, Department of Economics. Email: blackb@fas.harvard.edu. I would like to thank Mariana Colacelli, David Evans, Julie Mortimer, Ariel Pakes, Bryce Ward, and participants at the Harvard IO Workshop for helpful suggestions. I stake sole claim to any remaining errors.

> "Illegal downloading of music is theft, pure and simple. It robs songwriters, artists, and the industry that supports them of their property and their livelihood. Ironically, those who steal music are stealing the future creativity they so passionately crave. We must end the destructive cycle now."
>
> *-Frances W. Preston, President of BMI*[1]

1 Introduction

The short-run effects of file sharing on the sale of recorded music has been studied by several authors, as well as in the popular press and the courtrooms.[2] However, the short-run (static) welfare implications of file sharing are straightforward. Recorded music is priced well above marginal cost, and file sharing technologies such as Napster or Kazaa provide methods of transmission at marginal cost. Therefore, conditional on the existence of the good, file sharing is welfare-increasing for society as a whole in a static world. Of course, the world is not static, and therefore to understand the true impacts of file sharing on the market for recorded music, it is necessary to analyze how the advent of file sharing has affected firms' long-run supply decisions; that is, how has file sharing affected investment into the production of recorded music.

So far, the majority of the debate on file sharing has been on the simple, short-run question of how file sharing relates to demand:[3] "does file sharing increase sales or decrease sales for an album?," *conditional on the existence of the album*. As the above quote from Frances Preston indicates, the industry is concerned that if file sharing reduces the sales of albums, there may be reduced incentives to develop new music or "future creativity." This possibility should concern not only those that make their living in the recorded music

---

[1]RIAA 2003

[2]See, for example, Blackburn (2004), Oberholzer-Gee and Strumpf (2004), Rob and Waldfogel (2005), and Zentner (2004) among others.

[3]An exception to this is Mortimer and Sorenson (2005), which examines long-run trends in sales, release patterns, and concert activity in the music industry as a result of a change towards digital distribution.

industry, but society in general. The static welfare gains from acquiring music at marginal cost through unauthorized copying may well be eliminated if in the long-run, these dynamic effects reduce the supply of music. In this sense, the debate about how file sharing affects the recorded music industry in the long run, and how it may affect the film or television industries as well, is very similar to the literature on how patents and copyright affect research and development in general.[4]

Thus, in this paper, we draw on the existing literature on how file sharing affects the demand for albums in the short run in order to model and analyze the effect that this change has on recording company's incentives to invest in the production of albums. The implications of this model are then tested with data from 14 years of album sales charts to determine how file sharing has affected the development of new music and the ability of firms in the industry to bring that music to the market. Using the changes in the distribution of sales identified in Blackburn (2004), the model developed in this paper predicts that while incentives to produce albums from already established artists are reduced, the incentives to develop new music from new artists are increased. Although data limitations preclude definitive findings regarding these implications, the empirical results indicate that there has been little change in the patterns of album development and release since the development of file sharing technologies in 1999.

2 File Sharing and the Recorded Music Industry

We begin by providing a short history of the recorded music industry and how unauthorized copying through file sharing has developed and affected the industry. We start with a quick

[4]There is a tremendous amount of work done on the effects of patents and copyrights, in many different forums, so that a definitive summary would be beyond the scope of this paper. However, Lanjouw (1998) and Scherer and Weisburt (1995) have empirically investigated the impacts of patents on pharmaceuticals in India and Italy, respectively.

discussion of how the industry is vertically structured.

Albums are typically produced in the following manner. First, an artist, who is represented by a manager, is signed to multi-year recording contract by a record label, with compensation consisting of an up-front payment and then royalties from the sales of albums, generally between 5% and 13% of the retail price of the album (Standard and Poor's 2002). An album is then produced in one of the label's recording studios, printed onto a compact disc by the production arm of the owner recording company, and distributed by the distribution arm of said company. Thus, in addition to tight horizontal concentration, the path from artist to consumer is essentially completely vertically integrated. The typical distribution cost to retailers of an album hovers around $10 and a baseline industry figure is that the record company makes somewhere on the order of $5 per album sold (Billboard 2000), depending on the album specifics.

Meanwhile, distribution channels have also changed greatly since file sharing and the internet started to cause changes in the industry. When Napster was first introduced in 1999, 51% of albums were sold in retail music stores and 34% in "other stores." By 2002 and 2003 the share of sales in music stores had dropped to approximately 35%, with over 50% sold in "other stores" (RIAA 2004). Additionally, by 2003, fully 5% of all music sales occurred through the internet, a figure that has continued to grow (RIAA 2004). The general consensus in the industry is that this shift is a movement towards sales through large electronics chains such as Best Buy and Circuit City, as well as mass merchants such as Wal-Mart and away from small, localized music stores and chains.

It is also instructive to take a short diversion into the technology of file sharing and digital music files. Although the exact technology used to encode music into a digital format can vary, these differences are minor and so we focus on a discussion of MP3, or MPEG Audio Layer 3, files.[5] An MP3 file is a file which contains audio information to

[5] MPEG stands for "Moving Picture Expert Group."

played back on a computer or other electronic device (such as a MP3-capable CD player, or a portable MP3 player, such as an iPod). The MP3 format was developed in 1991 by the Fraunhofer Institute and uses what is called perceptual noise shaping to compress CD-quality audio tracks by a factor of 10-12, while providing almost the same level of sound quality (Bellis 2005 and Fraunhofer 2005). It is the compression technology that is the main advantage of digital audio files such as MP3 files; an uncompressed audio file requires approximately 10 megabytes of storage space per minutes of audio.[6] If the same audio file is compressed into an MP3, it would generally require less than 1 megabyte of space. Thus a typical song requires only 4 megabytes or so when stored as an MP3 file, but 40 megabytes when uncompressed.

This move to an advanced compression technology, coupled with typical internet access speeds in the late 1990s, and the introduction of file sharing networks allowed music consumers to copy (without legal authorization) MP3 files from remote computers at no cost other than the short amount of time required to download a 4 megabyte file. The first, and most famous, file sharing network was Napster, which was created in May 1999 by a Northeastern University student named Shawn Fanning. Napster became a huge success, with a reported user base of over 20 million unique user accounts worldwide at its peak, with routinely more than 500,000 unique IP addresses connected at any time (CNNMoney 2000). Up to the introduction of Napster, the recorded music industry in the United States was experiencing a huge period of growth. However, the gains made in the years prior to 1999 quickly disappeared and industry sources were quick to attribute this decline to the rapidly increasing popularity of Napster. As a result, the Recording Industry Association of America (RIAA) in December of 1999 filed suit in U.S. District Court (RIAA 1999) to have Napster dismantled. This began a long line of lawsuits which resulted in the end

[6]Thus, a blank 700MB compact disc holds approximately 70 minutes of uncompressed, standard CD-quality music.

of Napster in February 2001, although file sharing has continued on many other networks since then.

By October of 2002, the new king of file sharing networks was a program called Kazaa, which claims to have had nearly 400 million downloads of its software as of April 2005. Kazaa, and similar networks, proved harder for the RIAA to shut down, because the networks they use are decentralized, and utilize "peer-to-peer" technologies which mean that users are directly connected to each other rather than contacting each other through a central database, as was the case with Napster. The federal court system found that similarly decentralized networks (Grokster and Morpheus) were not liable for the unauthorized music trading taking place on their networks because they did not have direct control over what happens on them (CBS.com 2005). Therefore, in June 2003 the RIAA turned to a strategy of directly suing the users of file sharing networks in order to attempt to stop the unauthorized copying of digital music files. As of April 2005, the RIAA has continued to sue users of file sharing networks in groups of several hundred every few months. All of this legal action stems from a belief on the part of the music industry that the unauthorized copying of digital music files that occurs on file sharing networks is reducing the sales of recorded music. This point has been much contested in the popular press, the courtroom, and academic circles. The details of this discussion are explained below, but we will work under the assumption that file sharing has changed the distribution of sales in the recorded music industry, and seek to understand how this change is affecting the incentives for firms to invest in the creation of musical works.

A final change brought about by the emergence and importance of MP3 (and similar) technology is that the end of 2003 and the beginning of 2004 have seen the roll out of several new distribution channels utilizing legal MP3 downloads on a subscription, single track, or full album basis, starting with iTunes, and currently including offerings from Rhapsody, MusicMatch, Roxio's revamped Napster service, and even Walmart.com. By

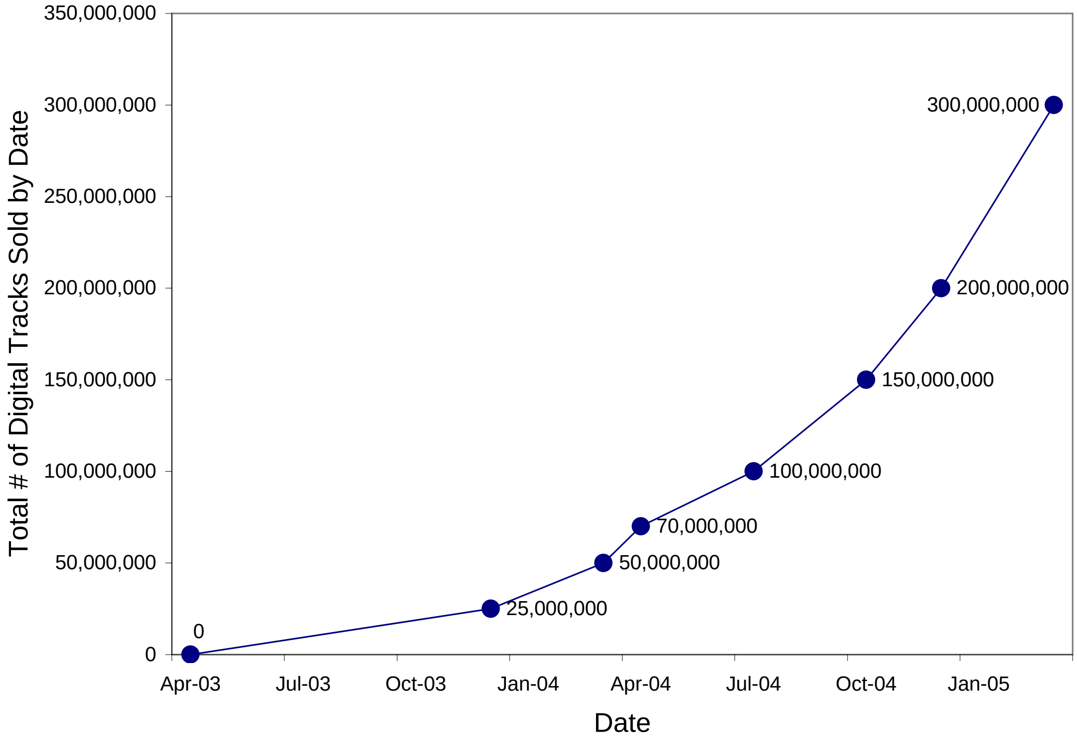

Figure 1: Cumulative Sales of Digital Tracks at iTunes.com, as reported at www.apple.com

far, the most successful of these enterprises has been iTunes, the music store owned by Apple Computer. According to press releases on the apple.com web site, the iTunes music store was opened on April 28, 2003, though at that time the store was only accessible to owners of Apple-brand home computers. It was not until October 16, 2003 that iTunes was made available to users of Microsoft Windows-based computers. iTunes generally offers for sale whole albums (which in this case means a bundle of individual digital audio files, but no CD, case, or cover art) at a price which is usually $9.99, or also individuals songs (a single digital audio file) for a price of $0.99. iTunes was successful very quickly, according to sales numbers released by Apple, again available at apple.com. Figure 1 characterizes the sales announcements from Apple; within two years of the initial release, Apple had sold over 200 million songs on iTunes, which given the pricing scheme, is approximately $200 million of revenue. This still pales in comparison to the size of the recorded music market, which was over $11 billion in 2004.

3 Review of Empirical Work on Effects of File Sharing

As mentioned above, almost all of the work on the effects of file sharing has focused on the short-run question of how file sharing affects the demand for recorded music. So far, the literature is not only small, due to both the newness of the phenomenon and the difficulty in obtaining reliable data on record sales and file sharing behavior, but also somewhat conflicted. There are two main sets of findings: First, several papers (see Hong (2004), Leibowitz (2004), Peitz and Waelbrook (2004), Rob and Waldfogel (2004), and Zentner (2003)), using either aggregates or survey data, find varying degrees of reduction in the demand for recorded music as a result of file sharing. Second, a recent paper by Oberholzer-Gee and Strumpf (2004) uses album-level data on sales and digital music file downloads and find no significant effect of file sharing on the demand for recorded music. The conflict here is striking, as the prima facia case against file sharing is very strong. When Napster is released in 1999, the RIAA reports that shipments of recorded music in the United States were approximately $15 billion, while in 2003 these shipments totaled only $11 billion.[7] Nevertheless, Oberholzer-Gee and Strumpf find no significant effect of file sharing on sales of recorded music.

Blackburn (2004) performs a similar analysis to that undertaken by Oberholzer-Gee and Strumpf, which reconciles the conflicts between these two sets of results. In particular, Blackburn finds that there are distributional impacts of file sharing wherein new artists (or relatively unknown artists) have the sales of their albums increased (through sampling, advertising, network effects, or the like) while established artists have the sales of their

[7] Both values are in real year 2000 dollars.

albums reduced, because there is less room for already well-known artists to gain market penetration than there is for new artists. Failure to account for these differences among artists causes one to find that for the average artist, file sharing has no impact on sales. However, there is both a real distributional impact as well as an aggregate impact, as established artists tend to sell more records than new artists.

It is this distributional impact that makes studying the investment incentive effects of file sharing interesting. Were it the case that file sharing reduces the sales for all albums, then the investment effects would be clear– the return to creating recorded music has dropped, and so there would be a reduced incentive to invest in the creation of such works. And, of course, if it was the case that file sharing has no impact on the demand for recorded music for any artist, then there would be no incentive effects. The fact that there is a change in the distribution of outcomes, and that file sharing has caused a shift in the market from established artists to unknown and new artists raises questions about the incentives the record companies face in deciding whether to develop new artists for the market or not. New artists now sell more when they start out, but if they become established artists, then their sales would be lower than in a world without file sharing. We now develop a theoretical model designed to highlight the effect of this shift in the distribution of sales and to understand the impacts it has on the incentive to bring recorded music to market.

4 The Model

The career of a music artist is represented as lasting two periods. In the first period, a recording company decides to sign a new artist or not and, if the artist is developed, releases a debut album. The debut album has some level of market success, and then the decision is made whether or not to develop the artist's career and release a follow-up album. The success of the follow-up album is dependent, in part, on the success of the initial album; that is, success breeds success. Inside this framework, we will introduce changes consistent with the effects of file sharing in order to analyze how file sharing has affected the incentives to invest in both new artists (first stage investment) and in existing artist development (second stage investment). Hendricks and Sorensen (2005) use a similar set-up to model a recording artists' career in examing how externalities between sales of different albums by the same artist affect record labels' incentive to invest in new albums from existing artists. They find that the size of these externalities are large enough to matter and lead to a form of "lock-in" whereby artists are essentially tied to their original label throughout their careers due to the fact that the original label internalizes the sales externalities across albums. We do not model competition among lables for an artist's contract and thus ignore these externalities in the model that follows.

The risk-neutral recording company is a monopolist with sole control over an artists fate; there is no competition for artists, and the decision to invest in an artist or not is the choice solely of the recording company. Further, we assume that sales from albums of different artists are independent of each other, and so the success of a particular artist is independent of the situation of other artists.

The career of artist i begins with a random draw of a level of talent m_i from an unspecified distribution F^m. The sales of a debut album, $s_{1,i}$, from an artist of talent m_i is taken (for analytical simplicity) to be exponentially distributed with mean βm_i. The parameter β will be used to measure the effect that file sharing has on the sales distribution of new artists. In order to produce a debut album, the recording company must pay a one-time fixed production cost of C, which is measured in units of albums sold.

After first period sales $s_{1,i}$ are realized, the second period begins. The recording company decides again whether to reinvest in artist i by paying a fixed cost K (again measured in units of albums sold) in order to develop and release a follow-up album. This set-up mirrors the standard recording contract, which stipulates that both parties agree to produce an initial album, and then the option to produce one or more follow-up albums (generally up to seven) is the sole decision of the recording label.[8] The follow-up album, if released, has sales $s_{2,i}$ which has expected sales (conditional on first period sales $s_{1,i}$) equal to $\gamma\tau s_{1,i}$. The parameters γ and τ are, respectively, the effect of file sharing on second period sales and a general scaling term which allows sales of follow-up albums to be, on average, higher or lower than sales of debut albums. No other assumption is made on the distribution of $s_{2,i}$ other than that it is distributed as a random variable with mean $\gamma\tau s_{1,i}$.

It is worthwhile to notice that second period sales are not directly affected by talent. Rather, first period talent is mixed with a random "market" process to determine the sales popularity of an artist, and second period sales are affected directly by this first period sales success. Therefore, in the model, firms do not make mistakes in the sense that the true distribution of sales is always known by the recording company. Thus, there is no process available by which file sharing can allow for more information about the underlying ability of an artist to achieve market success, as might be possible in the real world. Similarly, within this framework, it is impossible for an artist to improve over time, in the sense that the underlying talent level of the artist does not change over time, though it is, of course, possible, for an artist to sell more albums in the second period than in the first period. Introducing unobserved (to the record label) movements in the underlying level of the artist's talent would simply introduce noise into a (risk-neutral) problem; introducing observed movements in talent over time would change the decisions made by record labels, but should not change the effects of file sharing on those decisions.

All differences between artists in the model take the form of vertical differentiation in m_i, the talent of the artist (and ex post in $s_{1,i}$, the sales popularity of the artist). Were the model written with the heterogeneity taking the form of differences in costs (C and K) or career growths (τ), the exact form of the investment rules derived below would be different,

[8] See Passman (2003) for a detailed discussion of how music contracts typically work.

but the implications from changes in the amount of file sharing would be unchanged. If, however, the model were to allow heterogeneity among artists in how they are affected by file sharing (different β, γ, or f), then of course, the effects of file sharing would be determined by the precise nature of the assumptions about this form of heterogeneity.

This fully describes the model, and we can now solve the recording company's decision problem backwards, to determine the optimal investment rules in each period. Note that due to the monotonicity of the problem, investment rules take the familiar "cut-off" form of a lower bound for investment to be profitable.

4.1 Second Period Investment

The investment rule in the second period is straightforward. Conditional on first period sales, $s_{1,i}$, the expected profit (measured in albums sold) from developing a follow-up album is:

$$E(\pi_{2,i}|s_{1,i}) = \gamma\tau s_{1,i} - K$$

Thus, the recording company will invest in a follow-up album for all artists whose first period sales exceed:

$$s_1^* = \frac{K}{\gamma\tau}$$

4.2 First Period Investment

First period investment is slightly more complicated. The expected profits from a debut album are simply:

$$E\pi_{1,i} = \beta m_i - C$$

Ex ante expected second period profits must take into account the probability that a follow-up album is not made, and thus are:

$$\begin{aligned}
E\pi_{2,i} &= E(\pi_{2,i}|s_{1,i} > \frac{K}{\gamma\tau})\Pr(s_{1,i} > \frac{K}{\gamma\tau}) + E(\pi_{2,i}|s_{1,i} \leq \frac{K}{\gamma\tau})\Pr(s_{1,i} \leq \frac{K}{\gamma\tau}) \\
&= E(\pi_{2,i}|s_{1,i} > \frac{K}{\gamma\tau})\Pr(s_{1,i} > \frac{K}{\gamma\tau}) = (\gamma\tau(\beta m_i + \frac{K}{\gamma\tau}) - K)\exp(\frac{-K}{\gamma\tau\beta m_i}) \\
&= \gamma\tau\beta m_i \exp(\frac{-K}{\gamma\tau\beta m_i})
\end{aligned}$$

Therefore, total ex ante expected career profits for an artist of "talent" m_i are:

$$\begin{aligned}
E\pi_i &= E\pi_{1,i} + E\pi_{2,i} \\
&= \beta m_i - C + \gamma\tau\beta m_i \exp(\frac{-K}{\gamma\tau\beta m_i}) \\
&= m_i\beta(1 + \gamma\tau\exp(\frac{-K}{\gamma\tau\beta m_i})) - C
\end{aligned}$$

and therefore, the first period investment rule is to invest in any artist of talent $m_i > m^*$, where m^* is implicitly defined by:

$$\pi(m^*) = m^*\beta(1 + \gamma\tau\exp(\frac{-K}{\gamma\tau\beta m^*})) - C = 0$$

4.3 File Sharing

We now allow file sharing to affect the distribution of sales, by scaling the mean level of sales in each of the two periods of an artist's life. In particular, Blackburn (2004) finds that file sharing has had a distributional effect on sales- sales of recordings from new artists are increased through the existence of file sharing, while sales of recordings from established artists are reduced. Thus, letting f denote the extent of changes in sales due to file sharing, define:

$$\beta = 1 + f$$
$$\gamma = 1 - f$$

Thus, file sharing increases mean sales in the first period of an artist's life (when she is a new artist) and decreases it in the second period of an artist's life (when she is an established artist). The specification approximately models file sharing as proportionately raising sales for new artists as much as it reduces sales for established artists. This is also consistent with Blackburn (2004), which argues that not accounting for these distributional effects and using a data set with as many new artists as established artists results in finding an "average" file sharing effect of zero, as in Blackburn and also Oberholzer-Gee and Strumpf (2004).[9]

We can now examine the effect of an increase in the extent of file sharing on the investment rules. The effect of file sharing on the second period investment rule (whether or not to develop a follow-up artist) is straightforward, as first period sales have already occurred, and thus only the negative effect of file sharing on second period expected sales remains.

Proposition 1 *Increases in file sharing reduce the incentive for recording companies to invest in the creation of follow-up albums from established artists. In particular, increases in file sharing raise the minimum level of debut album sales necessary to make a follow-up album profitable. That is:*

$$\frac{\partial s_1^*}{\partial f} = \frac{\partial s_1^*}{\partial \gamma}\frac{\partial \gamma}{\partial f} = \frac{K}{\gamma^2 \tau} > 0$$

Proof. The proof is immediate and comes from simply differentiating the cut-off rule equation with respect to the file sharing variable, f. ■

[9] For algebraic clarity, we are making only a clean distinction between new and established artists, rather than allowing for degrees of knowledge about established artists. In the estimation that follows, we relax this assumption and control for the ex ante popularity of the artist as in Blackburn (2004).

The effect of file sharing on the first period investment rule is more complicated, as there are competing effects from file sharing on first period sales and second period sales, conditional on a follow-up album being created:

$$\begin{aligned}
\frac{\partial m^*}{\partial f} &= \frac{\partial m^*}{\partial \beta}\frac{\partial \beta}{\partial f} + \frac{\partial m^*}{\partial \gamma}\frac{\partial \gamma}{\partial f} \\
&= \frac{\partial m^*}{\partial \beta} - \frac{\partial m^*}{\partial \gamma} \\
&= \frac{-m^*}{\beta} + \frac{m^*}{\gamma}\frac{(K + \gamma\tau\beta m^*)}{\beta m^*(\exp(\frac{K}{\gamma\tau\beta m^*}) + \gamma\tau) + K} \\
&= \frac{-m^*}{\beta}(1 - \frac{\beta}{\gamma}\frac{1}{1+\theta})
\end{aligned}$$

where $\theta = \frac{m^* \exp(\frac{K}{\gamma\tau\beta m^*})}{\frac{K}{\beta} + \gamma\tau m^*}$, and the sign of $\frac{\partial m^*}{\partial f}$ depends on parameter values.

$$\frac{\partial m^*}{\partial f} < 0 \Leftrightarrow \gamma > \frac{\beta}{1+\theta}$$

Intuitively, the sign of $\frac{\partial m^*}{\partial f}$ depends on whether the negative effects of file sharing in the "established" period of an artist's life, γ, is overcome by the increase in career sales caused by the increased probability of higher second period sales as a result of a better performance from the debut album. That is, file sharing makes "star" artists sell fewer albums, but makes it easier for new artists to become stars. The question of whether or not it is better to invest in new artist is the question of whether the increased probability of becoming a "star" outweighs the reduction in sales from "star" artists. Thus, the effect that file sharing has on the incentive to invest in new artists is ambiguous in theory, and we will appeal to the empirical results in Blackburn (2004) in order to calibrate the parameters of the model to evaluate the derivative numerically.

Finally, before moving on to the calibration, we discuss how changes in cost parameters would affect investment, as it is likely that at the same time that technological advancements are making file sharing possible, similar changes are affecting the cost of producing and promoting albums. It is straightforward to see that increases in the cost of producing a debut album, C, raises the cut-off rule, m^*, and thus reduces investment in new artists, while increases in the cost of producing a second album, K, or a decrease in the size of follow-up album sales relative to debut album sales, τ, both raise the follow-up album cut-off rule, s_1^*, and thus reduce investment in established and new artists.

4.4 Calibration

Given that the sign of $\frac{\partial m^*}{\partial f}$ is ambiguous theoretically, we appeal to the results in Blackburn (2004) in order to calibrate the parameters of the model (f, τ) to actual data on the distribution of sales under file sharing, and a counter-factual world in which file sharing does not exist. In particular, f can be calculated as the percentage increase in sales for new artists when file sharing exists relative to when file sharing does not exist (or similarly, the percentage decrease in sales for established artists when file sharing exists relative to sales when file sharing does not exist).

First, in calibrating f, we see in Blackburn (2004) various estimates to the effect that file sharing has on sales for new artists and established stars. We take an estimate from counter-factual exercises in which 30% of file sharing activity is removed across the board. This estimate shows that an increase of 30% of file sharing activity increases sales of new artists by approximately 10% at the mean, while decreasing sales of albums from established stars by about 10% at the mean. Extrapolating this elasticity out, we proceed assuming that the total effect of file sharing is to increase sales of new artists by 33% at the mean and decrease sales of established artists by 33%.[10] We therefore set $f = 0.33$, so that $\beta = 1.33$ and $\gamma = 0.67$.

[10] As discussed in Blackburn (2004), this estimated elasticity is, for a variety of reasons, likely an upper-bound to the effect. Using other values of f between 0.1 and 0.33 does not change the results.

It is then left to calibrate τ, which is the ratio of mean second period sales to first period sales. Obtaining a calibration for this parameter is slightly more detailed. In principle, all that is needed is to take the ratio of sales between follow-up and debut albums. However, follow-up albums only appear if the debut album is successful enough for the the record company to decide to release a follow-up album. Thus, since we observe mean follow-up album sales only for artists who have a follow-up album released, we must take the mean of debut album sales only for artists who have (or will have) a follow-up album released. This becomes particularly difficult, both because the sales data in Blackburn (2004) contain only one album per artist and also because we can not predict with certainty which albums will have a follow-up released.[11] Therefore, a value for τ is constructed as follows:

- Determine what proportion of new artists have a follow-up album released and call it x;
- Calculate mean sales of debut albums conditional on sales being above the $1 - x$ percentile of debut sales and call it $\overline{s_1}$;
- Calculate mean sales of follow-up albums and call it $\overline{s_2}$;
- Use $\frac{\overline{s_2}}{\overline{s_1}}$ as an estimate of the ratio of mean sales of follow-up albums to debut albums;
- Finally, because second period sales may incorporate sales from multiple albums, calculate the average number of follow-up albums per artist with a follow-up album, and call it n;

[11] Worse, as discussed later in the Data section, we do not observe actual releases, but only releases that (a) are successful enough to appear on the Hot 200 Sales Chart for at least one week and (b) have been released by the end of the dataset. In this case, while we find that 47% of debut albums have a follow-up, we are underestimating both the number of debut albums as well as the number of follow-up albums. While there is no evidence that suggests either is more biased, it would seem to be the case that a debut album is more likely to be released and not appear on the chart than a follow-up album is. Thus, if anything, the 47% is an over-estimate, which would lead to an under-estimate of mean debut album sales, and thus an over-estimate of τ.

- Then use $n\frac{\overline{s_2}}{\overline{s_1}}$ as an estimate for τ.

First, we note (using the Billboard Hot 200 Chart data described below) that 47% of artists who have an album appear on the chart have a second album appear on the chart.[12] We then turn to the data from Blackburn (2004), which has U.S.-level sales data for a sample of 197 albums during the end of 2002 and most of 2003. This sample has 78 albums classified as being from new artists, and 119 classified as follow-up albums. See Blackburn (2004) for more details on the data. We use this data to calculate that $\overline{s_1} = 709,552$, that $\overline{s_2} = 524,237$, and that $n = 3.5$.[13] With these estimates, we can take an estimate of τ to be $\tau = 3.5\frac{524{,}237}{709552} \approx 2.6$.

It should be noted that we have no observable measures of the costs involved with producing (and, presumably, promoting) an album, either for a new artist (which is parameterized as C) or for an established artist (which is parameterized as K). We later discuss what effects changes in the parameters would have with respect to our empirical implications and tests thereof, but for now note that only K enters into the partial derivative $\frac{\partial m^*}{\partial f}$ directly, though C enters indirectly though its effect on m^*.

For any possible values of K and C, given the parameter values for f and τ above, m^* can be calculated from $\pi(m^*) = 0$, and thus the partial derivative $\frac{\partial m^*}{\partial f}$ can be calculated. First, assuming that $K = C$, we find that for any reasonable value of K (below $1,500,000$), the partial derivative is negative, which tells us that increases in file sharing serve to reduce the cut-off "talent" level for the development of a debut album. In order to make the sign of the effect negative, it must be that C greatly exceeds K. Interpreting the second period of an artist's career to include the sales from multiple album, as was done in the calibration

[12] Here we are appealing to the model's assumption that record labels choose to produce follow-up albums only based on the sales level of the first album. In the data, the raw correlation between the peak sales position of a debut album and observing a follow-up album on the Hot 200 chart is 0.25. While this is not overwhelming, it should be noted that (peak) sales positions do not perfectly correlate with (total) sales, and observing a debut/follow-up album is not the same as one existing.

[13] Again, the parameter n is underestimated both because some follow-up albums are not observed, and due to the right-censoring of the data. (That is, some follow-up albums are yet to be released.)

above, implies that this is very likely the case. If K is interpreted as the cost of producing and promoting multiple follow-up albums while C is simply the cost of producing and promoting the intial debut album, then K is surely larger than C.[14] Thus, the calibration exercise shows that file sharing causes firms to increase investment in new artists, as the gain in expected debut album sales and the increased probability of becoming a "star" and having relatively high follow-up album sales outweigh the loss (on average) in follow-up album sales.

5 Empirical Implications and Tests

5.1 Implications

The model above gives two clear predictions for the evolution of investment in the music industry as a result of file sharing. Conditional on the modeling assumptions, the first implication is unambiguous theoretically, while the second comes from the calibration exercise performed above. In particular, comparing the recorded music industry before and after file sharing, we expect to see:

1. In order for a recording company to develop an existing artist's career, and invest in a follow-up album, an artist's debut album must sell more copies in the face of file sharing than before file sharing existed.

2. After the introduction of file sharing, there should be more albums from new artists, as file sharing has increased the return to investing in new artists. This is further highlighted by the fact that file sharing reduces the return to investment in existing

[14]There is also some evidence that albums are more expensive to produce for more established artists than for new artists. On page 19 of *This Business of Music*, Krasilovsky and Shemel (2003) argue that production costs for new artists range from \$80,000-\$150,000, while they can approach or exceed \$500,000 for star artists. Without hard data on marketing costs, one might imagine that a similar relationship exists for those as well.

artists, and so new artists have an even better return relative to that from investing in existing artists.

These are clean implications that would ideally be simple to test empirically. However, due to data limitations, clear tests of these implications are not possible. As described in the next section, data are only available for albums which appear on the Billboard Hot 200 Chart at some time between November 7, 1992 and April 2, 2005, meaning that we only observe an album release if it is successful enough of an album to rank in the top 200 of album sales nationally at least one week.[15] While this is typically not too hard for a major label release, it is possible and even likely that there are a lot of changes in the investment and release patterns of artists affiliated with smaller, independent labels (or even self-producing and self-promoting artists). Biases caused by this data selection problem will be discussed as appropriate when detailing the empirical tests.

On the other hand, albums which do not ever sell enough to reach the top 200 in a single week are albums which have (by definition) minimal impact on the market and thus, presumably, social welfare. The exception, of course, is an artist who gets a release on a small label and proceeds to sell enough copies to convince a major label to sign her and develop a follow-up album, which may be successful enough to appear on the charts. In this case, our data will treat this artist as "debuting" on the major label, and will not see the initial, small-label release. To the extent that this serves as a new channel for major labels to identify and recruit artists, we are understating the total amount of investment in new artists.

[15] Using data on the number of new albums released each year between 1993 and 1999 obtained from the RIAA's 1999 year-end marketing report, approximately 2% of all new album releases reach the top 200 sales chart. However, back of the envelope calculations using the regression results from Table 6 suggest that these 2% of all releases account for over 95% of all sales.

5.2 Data

The primary data source for the empirical analysis comes from Billboard magazine. Each week, using data collected by Neilsen/SoundScan, Billboard magazine publishes the Hot 200 Album Sales chart, which ranks the top 200 selling albums in the United States in the past week. These weekly data are used to identify the existence of albums. With the data reported in the chart listings, we identify (for each album in the top 200 of sales that week) the artist that makes the album, the "release date" of the album, which is defined to be the first week that it appears on the Hot 200 chart, the peak position on the chart that it achieves, and the last week that the album appears on the chart.

Thus, we track artists' careers starting with the first album that ever appears on the Hot 200 chart and every subsequent follow-up album that appears on the chart, within the timeframe of the data sample.[16] Therefore, we observe the realization of all successful investments made by the recording companies in the industry. The chart data runs weekly from the week of November 7, 1992 through the week of April 2, 2005, for a total of 640 weeks of chart data. Additionally, chart data from the four years prior to the start of this data set (from November 12, 1988 until October 31, 1992) are used to establish artist histories for all artists who appear on the chart during the 640 weeks of primary data. For example, an artist who appears on the sales chart is designated as a new artist if she had not previously appeared on sales chart in any prior week, including the time period from November 12, 1988 until October 31, 1992. This four-year buffer is meant to distinguish (imperfectly) truly new artists from artists who have just gone a long time between album releases. Weekly sales chart were not made available prior to November 12, 1988 and then 1992 is chosen to be the start of the main sample to strike a balance between the chance of mislabeling an artist as new and providing equal time periods in the main sample both

[16]Of course, again, we do not see albums that may be produced and not released, or albums that are released but fail to ever appear on the sales chart.

before and after the arrival of Napster (which is released in May of 1999). Still, some artists will be miscategorized, and this will occur less as time passes in the sample. Thus, by construction, we are biased towards seeing too many "new" artists in the early years of the sample.

Finally, in order to focus on investment in the production of new music by recording companies, albums are only considered for the sample if they are of primarily new material, which eliminates live performance albums, greatest hits albums, and other compilation albums, such as the recent "Now That's What I Call Music" series. Additionally, movie and Broadway soundtrack albums (and any other albums which are credited to "Various Artists") are also excluded from the sample, for two reasons, as these albums are typically comprised of music from multiple artists, which makes it impractical to identify the artist responsible for the music on the album.

It should be noted that the Neilsen/SoundScan data from which the Billboard charts are built are point-of-sales data from retail and internet locations, from what is (according to Neilsen/SoundScan) approximately 90% of the retail market in the US, including both brick & mortar and internet points-of-sale. However, only physical album sales are included, so album (or single song) purchases from iTunes, Rhapsody, or other MP3 retailers are not included in the charts. What affect would including MP3 sales data have? It should first be noted that despite the large amount of attention paid to MP3 sales, they are still small in comparison to the sales of physical CDs. According to Billboard magazine, sales of physical albums were nearly five times (approximately 665 million to 135 million) sales of digital tracks in 2004 (Billboard 2004).[17] In order for the exclusion of MP3 sales data to impact the results, it would have to be that impacts of file sharing on release patterns differ for albums that sell more (or less) MP3s relative to physical album sales. In fact, this

[17] And, of course, each digital track is only one song, while an album sale bundles the purchase of many songs. Before 2004, sales of digital tracks were virtually non-existant.

is likely the case, as the sporadic data on MP3 sales and file sharing downloads seem to suggest that songs that are most likely to be downloaded on file sharing networks are also the most popular ones on legitimate MP3 services. Thus, the data employed in this paper is likely understating the sales of songs/albums for which file sharing is most important. However, this is unlikely to matter much in the estimation that follows, as the time period for which MP3 sales are even slightly important is a small fraction of the dataset, even in the post-Napster era. To the extent that it would matter, it is likely to bias against finding the implication above that debut albums will need to sell more in order for the record label to option a second album, as debut albums (which are among the types of albums most likely to be helped by file sharing, according to Blackburn (2004) and so the successful ones will also likely be the ones with relatively more file sharing activity) will have sales numbers in the data that are "too low," thus biasing the data against the empirical prediction from the model.

In addition to the sales chart data, data from the FCC's study of Local Telephone Competition and Broadband Deployment is used to identify the total number of installed broadband (high-speed) internet access lines in the United States. The FCC counts high-speed lines as connections that deliver services at speeds exceeding 200 kilobits per second (kbps) in at least one direction, and counts lines that connect both homes and business to the internet. These data are collected semi-annually, and linear interpolation is used to fill in the data for weeks in between the FCC data points. Lastly, there are several weeks in the data sample before the first FCC data point and after the last FCC data point, for which no data are available. A flexible polynomial specification is used to predict the number of high-speed lines in these outlying weeks using the data points from the FCC. Only a quadratic specification is necessary, and yields an R-square of 0.999.[18] The out-of-sample

[18]This is high enough that one might wonder about the FCC's data collection methods. According to the semi-annual Local Telephone Competition and Broadband Deployment report from which the data is taken, the FCC collects data "from providers with at least 250 highspeed lines in a state" and the data is taken from

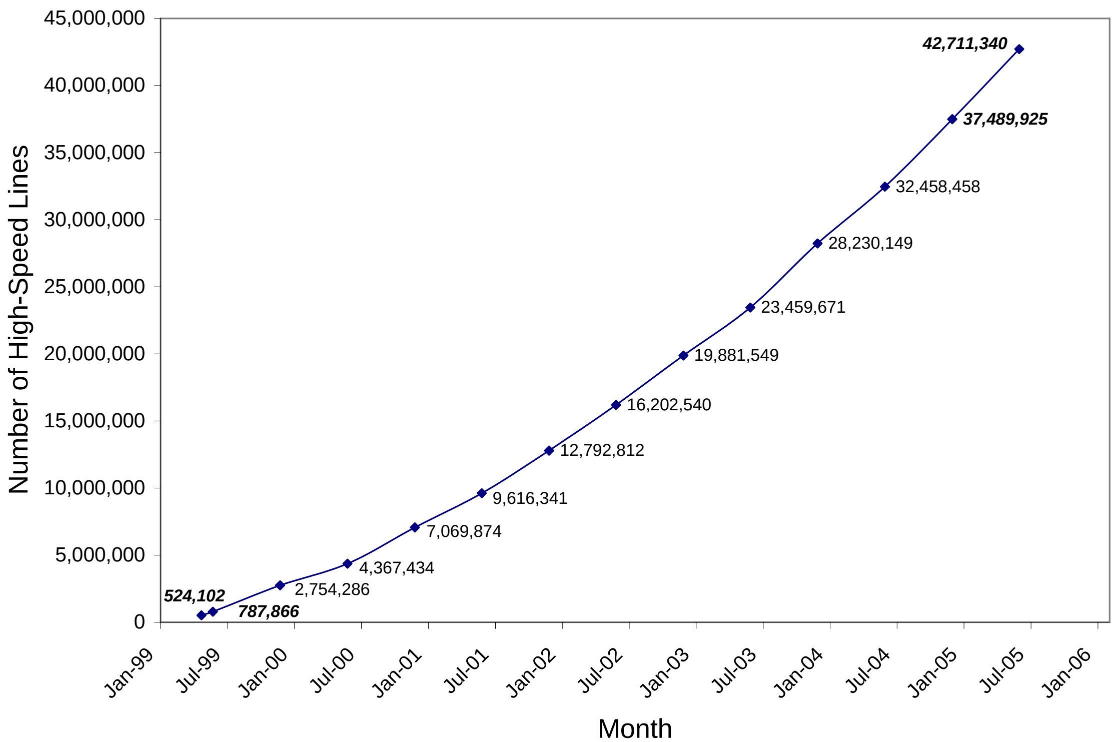

Figure 2: Number of Installed High-Speed Internet Access Lines in U.S., from the FCC

predictions from the specification are used to fill in the 31 weeks prior to the first FCC data point (and after the release of Napster) and the 43 weeks after the last FCC data point. Figure 2 plots this data; the standard font numbers are the actual FCC data points, and the bolded and italicized numbers are the predicted values from the quadratic specification.

We now provide a quick summary of the variables in the data, starting with a look at the album-level data. Table 1 presents summary statistics for the 6,454 albums of new material from a single artist that appear on the chart during the sample period. Fully 30% of albums are debut albums from new artists, where again, a new artist is one who, going back to November 1988, has not appeared on the chart previously. As for the market success of these albums, the average album stays on the chart for 18.8 weeks, and reaches a peak

---

"FCC Form 477 in the Commission's local competition and broadband data gathering program."

Table 1: Summary Statistics

Album-Level Data				N=6454
	Mean	St. Dev	Min	Max
New Artist	0.30	0.46	0	1
Length of Time on Chart	18.8	25.0	1	389
Peak Position (Inverted)	123.2	61.0	1	200
Artist-Level Data				N=2542
	Mean	St. Dev	Min	Max
New Artist	0.77	0.42	0	1
Length of Time on Chart	35.7	61.8	1	676
Peak Position (Inverted)	127.6	61.5	1	200
Number of Albums	2.99	2.88	1	32
New-Artist Level Data				N=1967
	Mean	St. Dev	Min	Max
Follow-Up Album Exists	0.47	0.50	0	1
Peak Position (Inverted)	106.1	61.2	1	200
Length of Time on Chart	17.2	23.8	1	148
Week-Level Data				N=640
	Mean	St. Dev	Min	Max
# of New Albums on Chart	10.1	5.8	1	31
# of New Albums from New Artist (Debut Albums)	3.1	2.1	0	11
After Napster Released	0.48	0.50	0	1
# of High-Speed Internet Lines	17,745,543	12,353,664	524,102	40,903,927

position of 77 on the chart. The variable peak position is inverted so that higher numbers correspond to higher positions on the chart. Thus, an album which reaches #1 on the Hot 200 chart will be assigned a "peak position" of 200, so the mean of 123 for peak position means that the average album reached a peak of #77 on the chart. Again, the right censoring of the data should be noted; at the end of the data sample there are 200 albums that are still on the charts (and maybe some more that may return) which may have stayed on the charts longer, or reached a higher peak in the future.

The second part of Table 1 provides a look at the artist-level data. There are 2,542 artists who appear in the main data sample, and 1,967 of those (77%) are classified as new artists. The average artist reaches a peak of #72 on the Hot 200 chart over the course of her career, and has an average of 3 albums on the chart over the data sample (though the median artist has only 2 albums on the chart). The average artist appears on the chart for

36 weeks over the sample period, though this is again highly skewed as the median artist appears on the chart for only 14 weeks.[19] Looking only at "new artists," of which there are 1,967, we see that 47% of new artists have a follow-up album released before the end of the sample period. Debut albums (the first album released by a "new artist") have an average peak of only #94 on the chart, and last for an average of only 17 weeks, with a median of 7.

Finally, the last part of Table 1 provides a look at the data at the week level. Here, we see that, on average, there are 10 new albums that enter on the sales chart each week, and that 3 of those, on average, are from new artists. Out of the 640 weeks in the data sample, 48% are after the release of Napster, and, looking at those weeks, we see that on average there were 17,745,543 broadband lines installed in the United States.

5.3 Testing the Empirical Implications

5.3.1 Results

We now turn towards testing the implications concerning the investment in musical talent that are derived from the model. Ideally, we would have complete data on all albums released over time, and would be able to track the composition of the releases among new artists and established artists, as well as the number of releases of both types of albums. Unfortunately, this data are not available without great expense.[20] Instead, as mentioned above, the available data that is on the appearance of albums in the Hot 200 sales chart over time. These data, then, create two main challenges. First, this is an ordinal ranking, rather than a cardinal one, and thus the model, which formulates the second period investment

[19]Tim McGraw appears on the chart for 676 weeks. As there are only 640 weeks in the data sample, this is an impressive feat. It means that, on average, there is more than one Tim McGraw album on the Hot 200 sales chart each week. Garth Brooks is second with 617 weeks on the chart.

[20]Purchasing a complete subscription to Neilsen/SoundScan's sales database would give a complete list of albums that were released and registered with them. Essentially all albums that are sold in stores somewhere in the U.S. are registered with SoundScan.

Table 2: Marginal Effects From a Probit on Album Being a Debut

	(1)	(2)	(3)
	Marginal Effects for a Probit on a Debut Being From a New Artist		
Time	**-0.00012**	-0.0001	-0.00011
	[0.00006]**	[0.00006]	[0.00009]
Napster Dummy	0.00198		
	[0.02339]		
# of BroadBand Lines		-0.0003	-0.00029
(Millions)		[0.00092]	[0.00092]
Log of # of BB Lines			0.00021
			[0.00158]
Observations	6454	6454	6454

Standard errors in brackets
* significant at 10%; ** significant at 5%; *** significant at 1%

cut-off rule in terms of total sales does not directly match up with data on ordinal rankings. Second, the required volume of sales in a given week to appear at different positions on the chart changes from week to week and over time in general, as more or less albums are sold overall. Moreover, the fact that we see only successful releases, rather than all releases, will bias our results, as discussed below.

We start by looking at the impact of file sharing on the composition of artists on the sales chart. If the implications from the model hold, then after file sharing develops, the sales charts should be composed of more debut albums from new artists than follow-up albums from established artists. There are two reasons for this, however. The first is the fact that new artists sell more albums relative to pre-file sharing days, while established artists sell fewer. The second is the incentive effects on investment— file sharing implies that we should see relatively more albums from new artists being released.

Table 2 provides results from a probit analysis on an album entering the chart being from a new artist.[21] There are few controls available here (since any historical artist characteristic immediately implies that an artist is not new), so we can simply allow for a time trend and look at the impact of a dummy variable indicating that Napster has been created,

[21]Throughout the paper, all probit results are reported as marginal effects, evaluated at the mean of all explanatory variables. For discrete variables, the reported marginal effect is the increase in probability for a discrete jump from zero to one.

or a variable measuring the number of installed high-speed internet lines.[22] Here we see weak evidence that over time there are fewer new artists appearing on the chart, but no evidence whatsoever that file sharing has affected this overall trend in any way.[23] Thus, we see no evidence of a change in the probability that an album on the chart is from a new artist after the development of Napster, despite two reasons to expect to. One possibility is that the change in sales towards new artists manifests itself through new artists that make the chart staying on the chart for longer, rather than more new artists making the chart. Still, though, the changing incentives should result in more new artists appearing on the chart; it is likely, then, that what is happening is that the predicted increased investment in new artists is resulting in a greater number of new artists who do not appear on the chart. This would, in fact, be consistent with the model in which artists are vertically differentiated (coming from an unspecified distribution of talent) and thus the increase in investment means producing albums from artists who are "worse" and thus more likely not to show up on the sales chart. The shortcoming of not having data for relatively poorly selling albums is biasing this test against being able to find this shift.

The most direct implication from the model is that in a world in which file sharing reduces the sales of established artists, record companies will require that in order to invest in releasing a follow-up album, debut albums must sell more copies. First, note that in this set of regressions, the dummy variable indicating the "Napster era" is lagged by one year. This is done because we are looking now at the outcome (appearance of a follow-up album on the sales chart) of a record company's decision made earlier. Thus, during approximately the first year of Napster, albums that are appearing on the chart were "greenlit" in the pre-

[22]The results do not change if we use a Napster dummy that equals one for weeks that are more than one year, one and a half years, or two years after Napster debuts; these variables would be appropriate if it took time for the industry to adjust, as is likely. Given that the number of installed broadband lines (which is increasing over time and would pick up a "delayed" reaction) is insignificant, this is not surprising. This is true throughout all of the analyses in the paper.

[23]Of course, as discussed above, seeing fewer "new" artists each week is likely due to the fact that over time, fewer and fewer established artists are potentially misclassified as new artists.

Table 3: Marginal Effects From a Probit on Having a Follow-Up Album Within Two Years

	(1)	(2)	(3)	(4)
	Marginal Effect (Hot 200 Debuts)		Marginal Effect (Hot 200 and Heatseeker Debuts)	
Time	-1.06E-05	5.79E-05	1.00E-05	-1.10E-04
	[1.137E-05]	[0.000129]	[6.92E-05]	[9.09E-05]
Peak Position on Chart (Inverted)	4.06E-04	4.36E-04	**-4.60E-04**	**-4.40E-04**
	[0.000305]	[0.000312]	**[0.00022]****	**[0.00023]***
Peak Position * 1-Year Lagged Napster Dummy	**-0.00107**		0.00013	
	[0.000547]**		[0.00027]	
Peak Position * # of BB Lines (In millions)		**-7.55E-05**		-1.37E-07
		[3.77E-05]**		[1.54E-05]
# of Weeks Spent on Chart	**0.012**	**0.012**	**0.00159**	**0.00156**
	[0.00256]***	**[0.00256]*****	**[0.00053]*****	**[0.00055]*****
# of Weeks on Chart * 1-Year Lagged Napster Dummy	0.00228		0.00101	
	[0.00146]		[0.00080]	
# of Weeks on Chart * # of BB Lines (In millions)		0.00011		0.00005
		[9.59E-05]		[0.00004]
Peak Position*Number of Weeks	**-4.76E-05**	**-4.71E-05**	1.18E-05	1.08E-05
	[1.38E-05]***	**[1.38E-05]*****	[1.05E-05]	[1.05E-05]
1-Year Lagged Napster Dummy	**0.136**		-0.011	
	[0.0663]**		[0.03153]	
Number of Installed BroadBand Lines (In millions)		0.0077		0.0028
		[0.00468]		[0.00224]
N	1591	1591	2586	2586

Napster era. Though the exact time from the record label's decision to the release of the album varies, conventional wisdom is that this lag is on the order of one year.[24]

Again, the data fails to provide sufficient detail for a perfect test, because for each week it reveals only the ordinal ranking of albums, and thus we can not search for an increase in an absolute cut-off rule for investment. And because file sharing reduces the overall sales distribution, the sales levels that the ordinal rankings correspond to are changing over time. To that end, we analyze the changing relationship between sales levels and chart positions below. For now, however, we investigate the relationship between the peak sales position obtained by debut albums as well as the length of time it remained on the sales chart and the probability that a follow-up album is released and appears on the sales chart. We include in the analysis only albums which debut within two years of the end of the data set, and run a probit analysis for the probability of a follow-up album being released

[24]Using no lag and using a one and a half year lag does not qualitatively change the results.

within two years, to avoid the right-censoring problem.[25] Table 3 provides results. The first two columns present results using the data described in the previous section. Columns (3) and (4) augment that data by using debut album data collected from Billboard magazine's Top Heatseekers chart, which tracks albums from "new and developing" artists that do not appear in the top 100 positions in the Hot 200 chart. This data are only used to identify debut albums, and therefore follow-up albums are taken only from the Hot 200 sales chart.[26]

Once again, the evidence is somewhat muddled. However, according to the marginal effects calculated in column (1) of Table 3, after the introduction of Napster, albums that peak higher on the chart gain less of an advantage in terms of the likelihood of the release of a follow-up album. In fact, comparing two debut albums similar in all other respects, if one album has a peak 100 positions higher than the other, that album is 10 percentage points less likely to have a follow-up album (relative to the other album) in the post-Napster world. This finding also holds when high-speed lines are used to proxy for file sharing use, as in column (2). However, there is no significant effect of reaching a higher position on the chart in either specification, and additionally, these effects go away when we expand the universe of debut albums to include those from the Heatseekers chart. Furthermore, the fact that we do not observe sales levels, but rather chart positions, tempers these results even further, as the level of sales associated with a top chart position is likely going down over time. Thus, even if record companies had not changed their cut-off rule, we would expect to see a negative coefficient on the interaction between the existence of file sharing and the peak chart position of an album.

On the other hand, we do see that the longer an album stays on the chart, the more

[25] A cut-off of two years is rather arbitrary. However, running the same specification with cut-offs of one and a half, two and a half, or three years yields the same qualitative results. When the specification is run as a probit on having a follow-up album, with no cut-off to attempt to work around the right-censoring problem, it becomes impossible to disentangle the effect of file sharing from the effect of the sample ending as both the napster dummy and the trend of broadband adoption are confounded with the right-censoring problem.

[26] Debut albums which only appear on the Heatseekers chart are assigned an (inverted) peak Hot 200 chart position of zero.

Table 4: Summary Statistics for # of Weeks on Chart for Debut Albums

Percentile	# of Weeks on Chart
1%	1
5%	1
10%	1
25%	2
50%	7
75%	22
90%	50
95%	71
99%	104
N	1967
Mean	17.2
S.D.	23.8

likely a follow-up album becomes. The estimates indicate that an album which stays on the chart for an additional week is 1.2 percentage points more likely to have a follow-up album released within two years. This is a strong effect; as can be seen in Table 4, the mean debut album appears on the chart for only 17 weeks, and the median debut album for only seven weeks. However, there is no evidence that this relationship is changing over time, either as a discrete jump upon the debut of file sharing technologies or gradually over time as file sharing becomes more prevalent through the expansion of broadband internet access.

A potential concern with this line of analysis is that as the music industry is changing due to the effects of file sharing, it's very possible that the amount of time that it takes to make an album is changing as well. This would cause problems with this sort of "cut-off" rule approach, as if the amount of time needed to produce an album is changing, then the relevent cut-off would also being changing over time. Mortimer and Sorensen (2004) argue that one logical response to the movement towards digital distribution (such as file sharing and legitimate MP3 downloads) would be artists and labels putting more effort into the production of albums. If this is so, then over time, the amount of time spent to produce an album would increase, and the relevent cut-off rule in the estimation above should be increasing as well. This would imply that in the Napster era, the cut-off rule used above is too short, and too few follow-up albums would be seen; that is, the estimated coefficients

Table 5: Chart Positions of Reported Sales Level Data

Chart Position	Frequency	Percentage of Total	Cumulative % of Total
1	328	24.0%	24.0%
2	256	18.7%	42.7%
3	141	10.3%	53.0%
4	89	6.5%	59.5%
5	52	3.8%	63.3%
6	52	3.8%	67.1%
7	45	3.3%	70.3%
8	35	2.6%	72.9%
9	35	2.6%	75.5%
10	26	1.9%	77.4%
11	26	1.9%	79.3%
12	14	1.0%	80.3%
13	18	1.3%	81.6%
14	17	1.2%	82.8%
15	20	1.5%	84.3%
> 15	215	15.7%	100.0%

for the Napster dummy and the # of Broadband Lines would be biased downward, which is fighting against the positive estimates presented in Table 3, and so we should not be too concerned about this potential bias.

5.3.2 The Relationship Between Chart Position and Sales

There is a limited amount of data available on how the relationship between sales levels and chart positions changes over time which might shed some light on this potential problem. Each week, in Billboard magazine, there is a column that is printed along the Hot 200 sales chart which discusses the week in music sales.[27] In a typical week, this column reveals the sales numbers corresponding to around four albums on the chart during the week. Going back through Billboard magazine to May 30, 1998 allows me to obtain weekly album sales and the corresponding chart position of the album for a sample of 1,369 album-weeks though April 2005. Of course, this sample of albums is in no way random. Typically, the albums that are discussed in the column are debut albums and the top sellers for the week. Table 5 reports the frequency with which different chart positions get the sales

[27]This column is currently titled "Over the Counter" and was titled "Behind the Bullets" before that. The column is typically written by Geoff Mayfield.

number reported in Billboard. Notice that 328 of the 1,369 observations are from the #1 album in the week, from a total of only 359 weeks.[28] While it is certainly the case that the albums that get reported are, by definition, newsworthy, it is not clear what the selection mechanism is. One might conjecture that abnormally high sellers would be more likely to be reported, but it is just as likely that abnormally low sellers would also get reported, and when reading through the articles both types of albums are frequently mentioned. Of course, the abnormally low sales of a #1 album is bounded below by the sales of the #2 album, while abnormally high sales are not bounded above. So it seems likely that there may be a upward bias in the reported sales of #1 albums, and to a lesser extent other very top albums, but it is still hard to determine how the bias might work for other chart positions. Unfortunately, we have no mechanism with which to determine what the selection process is and thus we have no ability to correct for it. The fact that we do not know the sign of the bias from this selection process is does not mean that there is no bias; however, we will attempt to use this data the best that we can, while acknowledging the fact that we are unsure of the biases involved. With the data selection issues in mind, we can examine how the relationship between (reported) sales positions and (reported) sales changes over time.

We first present Figure 3, which graphs of the average reported sales by category of reported chart position over time. Each bar in the chart represents the average sales of an album whose sales are reported in Billboard magazine in the corresponding range and year. We see that, in general, sales numbers seem to be lower for a given chart position in later years, though this is more pronounced for albums below the top ten than for albums in the top ten, which would be expected given the bias discussion above. Other than this general decline, there is no visible difference in the size of the effect for different chart positions.

[28] The #200 album is reported four times.

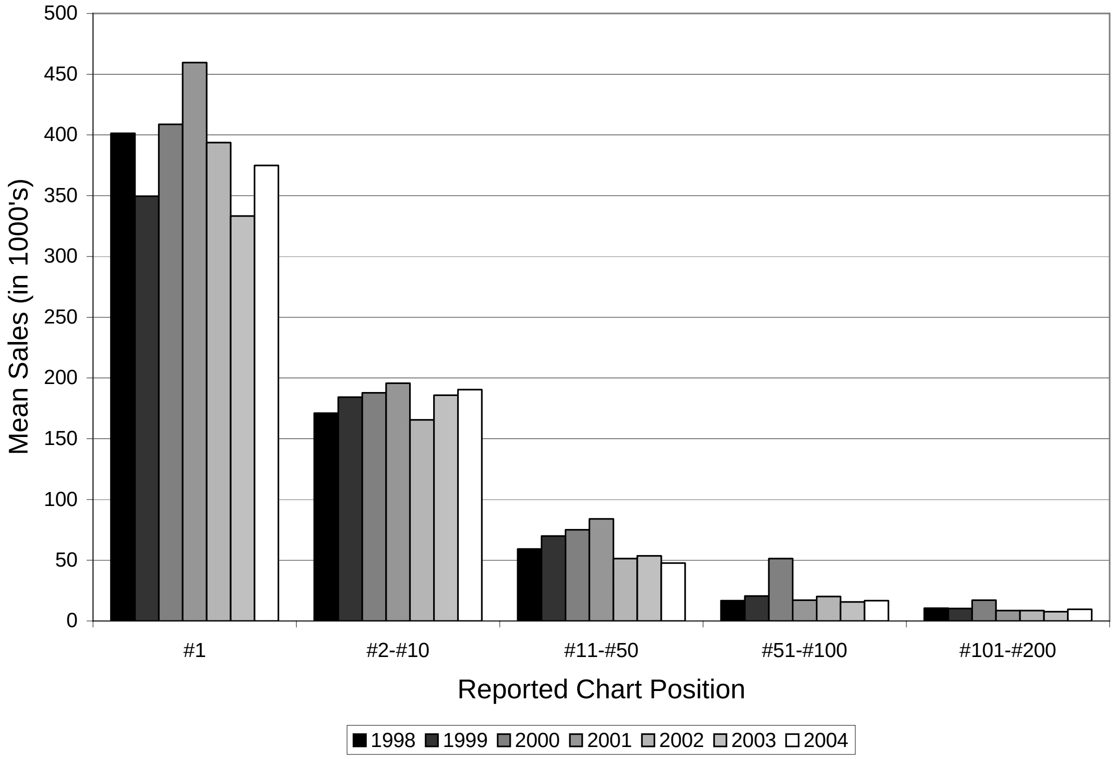

Figure 3: Mean Sales by Reported Chart Position

We then turn to the estimation of a simple regression of the following form:

$$\log(\text{sales}) = \alpha + \beta_1 \log(\text{chart position}) + \beta_2(\text{time trend}) + \beta_3 W + \epsilon$$

to determine if, over time, reported sales are falling after controlling for the chart position of the album. Due to the issues with the data from top selling albums, this is done both for all albums and only for albums below the top ten. The results of these regressions are given in Table 6. Column (1) presents the results for all albums, while Column (2) presents the results for only non-top ten albums. We see that, controlling for chart position (and holiday effects), there is a slight reduction in sales numbers, though that effect becomes statistically insignificant for non-top ten albums only. The evidence above, however, is

Table 6: Link Between Chart Position and Sales

	(1) Log of Sales (All Albums)	**(2)** Log of Sales (Non-Top 10)
Log of (Inverted) Chart Position	**1.30**	**0.64**
	[0.059]* **	**[0.045]*** **
Time Trend	**-0.0005**	-0.0004
	[0.0002] **	[0.0003]
Christmas Dummy	**0.796**	**1.23**
	[0.160]* **	**[0.39]*** **
# of Weeks Away From Christmas	**0.124**	**0.159**
(Up to Four)	**[0.018]*** **	**[0.029]*** **
Constant	**5.09**	**7.38**
	[0.21]* **	**[0.23]*** **
Observations	1369	284
R-squared	0.29	0.47

likely due to the reduction in sample size rather than the effect not existing for non-top ten albums. The coefficient of -0.0004, if correct, means that over the course of a year, mean sales (controlling for chart position) are reduced by approximately 2%.

Recall that the results in the above sub-section indicated that in the post-file sharing world, the probability of seeing the release of a follow-up album for a debut album that peaked 100 positions higher was 10 percentage points less than it would have been before file sharing. The analysis linking chart positions to sales levels indicates that we should attribute some of that change to the fact that sales numbers are generally lower in the post-Napster world. However, the magnitude of the drop in sales does not seem to be enough to explain the entire effect, as a change of 100 positions on the chart (for a peak) makes a tremendous difference in sales, and the 10% drop in the probability of a follow-up is nearly 20% of the mean probability, which is much more than would be expected from such a (relatively) small drop in chart position-sales over time. We can therefore attribute some of this change to a possible reduction in the willingness of recording companies to release follow-up albums for lesser-selling albums than in pre-Napster times, though the evidence is generally weak.

5.4 A Further Examination of Changes in Release Patterns

So, while we see only weak evidence supporting the predictions of the model, it is worth noting that there does not seem to be strong evidence that there is a dramatic shift in the distribution of albums that appear in the market. The first set of probit regressions suggested that there has not been a change in the proportion of albums on the chart that are from new artists relative to established artists, and the second set of regressions suggested that the advantages of selling more debut albums have diminished. While not providing resounding support for the model, the data does not provide any evidence of the industry's arguments that file sharing is eliminating the incentives to produce and release music to the market.

In fact, the coefficient on the "Napster" dummy in Column (1) of Table 3, which indicates that the debut album was released after the release of Napster, is large and positive, indicating that ceteris paribus a follow-up album is 14 percentage points more likely to be released within two years of the debut album after the release of Napster. Although not conclusive at this point, this evidence is not consistent with a world in which file sharing has reduced the incentives to supply new music. New artists and established artists fill the charts in the same proportion as in the past, but there seem to be more albums from established artists (and thus also more albums from new artists), which suggests that there is more turnover on the charts than in the past, but no less of a supply of albums. Of course, this effect goes away when looking at debut albums from the Heatseekers chart as well, which suggests that post-Napster, debut albums which are not succesful enough to reach the Hot 200 sales chart are less likely to produce a (succesful) follow-up than they were pre-Napster, as might be expected if file sharing increases sales of unknown artists. If an unknown artist can not sell records even with the help provided by file sharing, record labels would logically reach the conclusion that this artist is unlikely ever to be a money-maker.

We now turn to investigate these issues further by looking at the length of time that an album stays on the chart. In order to do this, I begin by using survival analysis techniques to measure what affects the length of time an album stays on the chart before it exits, never to return. I run two different specifications for modeling the hazard function, a Cox hazard and a Weibull hazard, which is free to exhibit either positive (more likely to leave as time passes) or negative duration dependence (less likely to leave as time passes).[29] The estimate of ρ captures this, with $\rho > 1$ indicating positive duration dependence. The results are presented in Table 7. Columns (1) and (2) present results from a Cox proportional hazard model, while columns (3) and (4) do the same using a Weibull hazard model. These regressions also use 52 dummies for the week of the year during which the initial album is released, as the time of year during which an album is released can help to control for unobservable characteristics of the artist (in the record label's eyes).[30] Also, from this point on, we are able to control for the ex ante popularity of the artist, as is done in Blackburn (2004). Thus, we define an ex ante popularity variable in exactly the same way; an artist's ex ante popularity for a given album is equal to 201 minus the highest Hot 200 Sales Chart position obtained in the past ten years, and for their debut album, artists are assigned an ex ante popularity of zero.

These first results indicate several notable relationships: First, album durations shorten over time, which is consistent with the finding that there are more debuts per week over time (as well as the above conjecture that turnover on the charts is increasing), but what is remarkable is that new artists tend to stay on the chart longer. The estimated hazard ratio of approximately 0.65 in all specifications means that a new artist is 65% as likely to exit the chart in a given week relative to an established artist. Looking at the effect of

[29]When using the Cox specification, the hazard is assumed to be: $h(t) = h_0(t)\exp(X\beta)$. The baseline hazard function, $h_0(t)$, is not estimated. In the Weibull specification, the hazard is assumed to be: $h(t) = \rho t^{\rho-1}\exp(X\beta)$. These assumptions about the hazard function are used throughout all survival analysis in the paper.

[30]In fact, the inclusion of the week-of-year dummies have very little effect on the magnitudes of the reported marginal effects, and never affect the sign or significance of them.

Table 7: Survival Analysis of Hot 200 Sales Chart Duration

	(1)	(2)	(3)	(4)
	Hazard Ratio for Time on Chart			
Time	**1.0013**	**1.0007**	**1.0014**	**1.0006**
	[0.0001]* **	**[0.0001]*** **	**[0.0001]*** **	**[0.0001]*** **
New Artist	**0.6687**	**0.6692**	**0.6236**	**0.6234**
	[0.029]* **	**[0.029]*** **	**[0.027]*** **	**[0.027]*** **
Ex Ante Popularity	1.0004	1.0003	1.0004	1.0004
	[0.0002]	[0.0003]	[0.0003]	[0.0003]
Peak Position of Album	**0.9841**	**0.9840**	**0.9823**	**0.9822**
	[0.0003]* **	**[0.0003]*** **	**[0.0003]*** **	**[0.0003]*** **
Total # of Debuts in Release Week	0.9975	0.9965	0.9973	0.9958
	[0.0034]	[0.0034]	[0.0035]	[0.0035]
Napster Dummy	0.9249		0.9203	
	[0.047]		[0.047]	
# of BroadBand Lines		**1.0071**		**1.0112**
(Millions)		**[0.0021]*** **		**[0.0022]*** **
Cox	Y	Y	N	N
Weibull	N	N	Y	Y
Week of Year Dummies	Y	Y	Y	Y
rho (Weibull parameter)			**1.0650**	**1.0700**
			[0.009]* **	**[0.009]*** **
N	6454	6454	6454	6454

file sharing, we see mixed evidence. The Napster dummy is insignificant, and the point estimates suggest that after file sharing, albums stay on the chart for a longer period of time, but the measures of broadband penetration suggest the opposite, and significantly. That is, as file sharing grows (through the growth of broadband penetration), albums stay on the chart for even shorter periods of time. Note that this is as much a demand effect as a supply effect, but it is noteworthy nevertheless.

We now turn to examining the same relationships, looking to see if the advent of file sharing affects new artists' longevity differentially than how it affects established artists. Given the results in Blackburn (2004), we may expect to see that new artists, whose sales are increased by file sharing, stay on the charts for even longer (relative to established artists' albums) after the introduction of file sharing. Table 8 displays the results of these regressions, once again with the first two columns using a Cox hazard function and the last two using a Weibull model.

Here, we see essentially the same set of results concerning chart duration, but there

Table 8: Survival Analysis of Hot 200 Sales Chart Duration, Differences by Artist Type

	(1)	(2)	(3)	(4)
	Hazard Ratio for Time on Chart			
Time	**1.0013**	**1.0007**	**1.0014**	**1.0006**
	[0.000150]* **	**[0.000143]*** **	**[0.000150]*** **	**[0.000142]*** **
New Artist	**0.6890**	**0.6744**	**0.6550**	**0.6363**
	[0.037]* **	**[0.0328]*** **	**[0.0348]*** **	**[0.0310]*** **
Ex Ante Popularity	1.0004	1.0004	1.0004	1.0004
	[0.000250]	[0.000250]	[0.000251]	[0.000251]
Peak Position of Album	**0.9841**	**0.9840**	**0.9823**	**0.9822**
	[0.000261]* **	**[0.000262]*** **	**[0.000260]*** **	**[0.000261]*** **
Total # of Debuts in Release Week	0.9973	0.9965	0.9971	0.9958
	[0.00344]	[0.00343]	[0.00346]	[0.00345]
Napster Dummy	0.9405		0.9460	
	[0.0507]		[0.0508]	
New Artist*Napster Dummy	0.9472		0.9147	
	[0.0531]		[0.0514]	
# of BroadBand Lines (Millions)		**1.0073**		**1.0118**
		[0.0022]* **		**[0.0023]*** **
New Artist*# of BroadBand Lines (Millions)		0.9992		0.9978
		[0.00240]		[0.00242]
Cox	Y	Y	N	N
Weibull	N	N	Y	Y
rho (Weibull parameter)			**1.0654**	**1.0702**
			[0.0094]* **	**[0.0094]*** **
N	6454	6454	6454	6454

is very little evidence that the chart durations of new artists and established artists are differentially affected by file sharing. In no specification is there any evidence that file sharing has affected new and established artists differently.

Finally, returning to the release patterns of albums, we look at the effect that file sharing has had on the timing of releases. It is impossible to observe changes in the amount of time spent developing debut albums, since there is no reference point to which we can compare the debut album's debut on the chart. Again we use survival analysis procedures, but in this context surviving though a time period means not releasing a next album. Table 9 presents the results, again in the form of hazard ratios and again with the first two columns displaying results from a Cox model and the last two displaying results from a Weibull model. It should be pointed out that in this specification, all albums are considered, and those for which a follow-up album does not exist, the model treats the observation as "surviving" throughout the entire sample period.

Table 9: Survival Analysis of Time Between Releases

	(1)	(2)	(3)	(4)
	Hazard Ratio for Release of Next Album			
Time	**1.0005**	1.0002	**1.0010**	**1.0015**
	[0.00020]**	[0.00018]	**[0.00020]*****	**[0.00018]*****
New Artist	**0.8736**	**0.8740**	**0.8485**	**0.8487**
	[0.0523]**	**[0.052]****	**[0.051]*****	**[0.051]*****
Ex Ante Popularity	**1.0039**	**1.0039**	**1.0040**	**1.0040**
	[0.0003]***	**[0.0003]*****	**[0.0003]*****	**[0.0003]*****
Peak Position of Album	**1.0026**	**1.0026**	**1.0035**	**1.0034**
	[0.00038]***	**[0.00038]*****	**[0.00038]*****	**[0.00038]*****
# of Weeks Spent on Chart	**1.0433**	**1.0437**	**1.0473**	**1.0471**
	[0.0046]***	**[0.0046]*****	**[0.0046]*****	**[0.0046]*****
Peak*# of Weeks on Chart	**0.9998**	**0.9998**	**0.9998**	**0.9998**
	[0.00002]***	**[0.00002]*****	**[0.00002]*****	**[0.00002]*****
Total # of Debuts in Release Week	1.0032	1.0021	1.0034	1.0043
	[0.0046]	[0.0045]	[0.0045]	[0.0046]
Napster Dummy	0.9969		1.1027	
	[0.062]		[0.070]	
# of BroadBand Lines		**1.0001**		**0.9932**
		[0.0034]***		**[0.0032]****
Cox	Y	Y	N	N
Weibull	N	N	Y	Y
rho (Weibull parameter)			**1.1507**	**1.1463**
			[0.015]***	**[0.015]*****
N	6454	6454	6454	6454

First, we note simple patterns: new artists spend more time before their second career album than established artists spend between album releases, as a follow-up album is 12% less likely to appear in a given week for a new artist than for established artists. Albums that are more successful either in time of total number of weeks spent on the chart, or the peak position obtained by the artist have less time before the release of the next album; record labels seem to try to “strike while the iron is hot” and release follow-up albums more quickly for artists that are currently successful. Unfortunately, we see that there is very mixed evidence on the effect of file sharing on the release time of albums. In these Cox specifications, file sharing, when measured as the total number of installed broadband lines, seems to decreases the time to a follow-up album, but in the Weibull specification, it increases the length of time between albums. This evidence is contradictory and sheds little light on the effect of file sharing on the time in between albums. We therefore instead turn to table 10, which displays results from similar analysis that allows file sharing to have

Table 10: Survival Analysis of Time Between Releases, Differences by Artist Type

	(1)	(2)	(3)	(4)
	Hazard Ratio for Release of Next Album			
Time	**1.0005**	1.0002	**1.0010**	**1.0015**
	[0.000203]* **	[0.000181]	**[0.000202]*** **	**[0.000181]*** **
New Artist	0.9216	0.9177	**0.8338**	**0.8462**
	[0.0628]	[0.0592]	**[0.0563]*** **	**[0.0539]** **
Ex Ante Popularity	**1.0039**	**1.0039**	**1.0040**	**1.0040**
	[0.00033]* **	**[0.00033]*** **	**[0.00033]*** **	**[0.00033]*** **
Peak Position of Album	**1.0026**	**1.0026**	**1.0035**	**1.0034**
	[0.00038]* **	**[0.00038]*** **	**[0.00038]*** **	**[0.00038]*** **
# of Weeks Spent on Chart	**1.0432**	**1.0434**	**1.0472**	**1.0471**
	[0.00461]* **	**[0.00461]*** **	**[0.00463]*** **	**[0.00464]*** **
Peak*# of Weeks on Chart	**0.9998**	**0.9998**	**0.9998**	**0.9998**
	[2.3E-05]* **	**[2.3E-05]*** **	**[2.3E-05]*** **	**[2.3E-05]*** **
Total # of Debuts in Release Week	1.0028	1.0018	1.0035	1.0044
	[0.0046]	[0.0046]	[0.0046]	[0.0046]
Napster Dummy	1.0261		1.0916	
	[0.067]		[0.072]	
New Artist*Napster	**0.8820**		1.0442	
	**[0.069]* **		[0.081]	
# of BroadBand Lines		**1.0104**		**0.9931**
(Millions)		**[0.0035]*** **		**[0.0033]** **
New Artist*# of BroadBand Lines		**0.9905**		1.0006
(Millions)		**[0.0048]** **		[0.0046]
Cox	Y	Y	N	N
Weibull	N	N	Y	Y
rho (Weibull parameter)			**1.1509**	**1.1464**
			[0.015]* **	**[.015]*** **
N	6454	6454	6454	6454

differential effects on new and established artists.

We again see the same contradiction between the analyses using Cox and Weibull specifications. Now, however, we see that in the Cox specifications, new artists take less time for the release of the follow-up album once Napster is released and as the use of broadband internet access grows. Unfortunately, this evidence is again mixed as a similar effect does not appear in the Weibull hazard models, and in fact is of opposite sign, though insignificant, in one of them. Thus it is again very hard to draw any conclusions from this data, but there is weak evidence that file sharing has reduced the time between albums for new artists.

In light of the failure of the data to reveal much support for the predictions of the model or the implications from Blackburn (2004) concerning differences in how file sharing affects new and established artists, it is prudent to discuss again the limitations of the available data. The predictions of the model regarding investments made by record companies are straightforward, but data that taken exclusively from the sales charts is not well suited for testing it. The set of albums that appear on the sales chart is merely a subset of all album releases, and a particularly biased one consisting only of relatively successful albums. As the predictions from the model concern decision rules relating the release of albums from the set of all potential new artists (which is not observed) and the release of follow-up albums from the set of all debut albums (which is also not observed), it is not too surprising that little is seen in the data.

That said, there is some information in the lack of information that the data provides. In particular, there is no evidence whatsoever of dramatic changes in the visible album release patterns by record labels. And this is in spite of the fact that there is lots of evidence that file sharing has caused dramatic changes in the patterns of albums sales. Were it the case that file sharing was severely hurting the incentives to develop new music (which is not what the simple model predicts), we would expect to have seen some changes in the number of new artists being successfully brought to market more than five years after file sharing was introduced. On the other hand, it is true that there have been other technological changes at the same time that file sharing has been growing which are likely decreasing the costs of producing records, and this effect (a reduction in K or C in the model) may mitigate the reduction in supply caused by file sharing. So, while we do not see any evidence of changes against the pre-Napster world, it may be that there are changes relative to the counter-factual world in which the cost-technologies have evolved, but file sharing did not. Unfortunately, the data at hand prevents conclusions regarding such conjectures from being definite.

6 Conclusions

This paper investigates how the changes in the recorded music industry due to the advancement of file sharing technologies has affected firms' incentives to develop new music, both from new and established artists. We developed a long-run model of investment in musical talent and calibrated it to the documented short-run changes to the distribution of sales. This methodology allowed us to examine the dynamic impacts on product availability. Our model predicts that while file sharing has made it less profitable to produce follow-up albums for established artists, effectively "raising the bar" of debut album success needed to make a follow-up album release worthwhile, file sharing has overall increased incentives for record labels to develop new artists. The reason for this finding is two-fold: first, sales of albums from new artists are disproportionately helped by file sharing and second, the increased visibility that arises from selling more copies of a debut album increases sales of future albums. These effects together outweigh the negative effect of the overall decline in sales of follow-up albums.

We then take the predictions of the model regarding changes in release patterns to data on album performance on the Billboard Hot 200 Sales Chart between 1992 and 2005. Although there is insufficient detail in the data to allow us to test the exact predictions of the model, the proxies that we employ indicate that file sharing has not drastically reduced the supply of albums to the market place. Precisely because the static welfare effects of file sharing are unambiguously positive, the lack of evidence against negative dynamic effects suggests that the societal harm of unauthorized copying in the recorded music industry is being overstated.

The findings of this paper, then, suggest that either file sharing is not reducing the incentives of recording companies to invest in the production of music, or that recording companies have already innovated in (unobserved) new technologies to ensure future benefits beyond what would be delivered by the previous business model focused primarily on the sales of recorded music.

References

[1] apple.com (2005), “Press Info: Recent Press Releases,” http://www.apple.com/pr/, downloaded on April 17, 2005.

[2] Bellis, M. (2005), “The History of MP3,” website, http://inventors.about.com/od/mstartinventions/a/MPThree.htm, downloaded on April 17, 2005.

[3] *Billboard Magainze* (1988-2005), “Chart Archives,” http://www.billboard.com.

[4] *Billboard Magainze* (2000), “Is Biz Poised For Renewed Price Wars?,” January 8, 2000.

[5] *Billboard Magainze* (2005), “Market Watch,” January 8, 2005.

[6] Blackburn, D. (2004), “On-line Piracy and Recorded Music Sales,” mimeo.

[7] CBS.com (2005), “Sharing Files And Filing Suit: A Timeline,” http://www.cbsnews.com/elements/2003/06/26/in_depth_scitech/timeline560590_0_main.shtml, downloaded April 17, 2005.

[8] CNNMoney (2000), "Napster: 20 million users," http://money.cnn.com/2000/07/19/technology/napster, July 19, 2000.

[9] Fraunhofer-Gesellschaft, “Basics about MPEG Perceptual Audio Coding: The purpose of audio compression,” http://www.iis.fraunhofer.de/amm/techinf/basics.html, downloaded on April 17, 2005.

[10] Hong, S. (2004), “The Effect of Digital Technology on the Sales of Copyrighted Goods: Evidence from Napster,” mimeo.

[11] Krasilovsky, M. and S. Shemel (2003), *This Business of Music: The Definitive Guide to the Music Industry*, 9th Edition, Watson-Guptil Publications.

[12] Lanjouw, J. (1998), “The Introduction of Pharmaceutical product Patents in India: ‘Heartless Exploitation of the Poor and Suffering’?,” NBER Working Paper #6366.

[13] Leibowitz, S. (2004), “File-Sharing: Creative Destruction or just Plain Destruction?,” mimeo.

[14] Mortimer, J.H. and A. Sorensen (2004), “Digital Distribution and Demand Complementarities: Evidence from Recorded Music and Live Performances,” mimeo.

[15] Oberholzer-Gee, F and K. Strumpf (2004), "The Effect of File Sharing on Record Sales: An Empirical Analysis." HBS mimeo.

[16] Passman, D. (2003), *All You Need to Know About the Music Business*, 5th Edition, Free Press.

[17] Peitz, M. and P. Waelbrook (2004), “The Effect of Internet Piracy on Music Sales: Cross-Section Evidence,” *Review of Economic Research on Copyright Issues*, Vol. 1(2), pp. 71-79.

[18] Recording Industry Association of America (1999), “Legal Cases,” http://www.riaa.com/news/filings/pdf/napster/Napster_Complaint.pdf

[19] Recording Industry Association of America (2004), “Consumer Purchasing Trends,” http://www.riaa.com/news/marketingdata/pdf/2003consumerprofile.pdf

[20] Rob, R. and J. Waldfogel (2004), “Piracy on the High C’s: Music Downloading, Sales Displacement, and Social Welfare in a Sample of College Students.” NBER Working Paper #10874.

[21] Scherer, F.M., and S. Weisburst (1995), "Economic Effects of Strengthening Pharmaceutical Patent Protection in Italy," *International Review of Industrial Property and Copyright Law*. No. 6, 1009-1024.

[22] Standard and Poor's (2002), "Movies And Home Entertainment," *Industry Surveys*, November 2002.

[23] Zentner, A. (2003), "Measuring the Effect of Online Music Piracy on Music Sales," mimeo.