An Empirical Study of Investment Externalities: The Case of Albums by the Same Recording Artist *

Ken Hendricks
Department of Economics
University of Texas at Austin

Alan Sorensen
Graduate School of Business
Stanford University

July 2005

Abstract

This paper studies the role of investment externalities in the relationships between record labels and recording artists, and the impact of these externalities on the structure of the market for recorded music. We show that the release of a new album increases sales of old albums by the same artist, and the increase is substantial and permanent—especially when the newly released album is a hit. The incumbent label (i.e., the owner of the artist's catalog) benefits from this externality and, more importantly, obtains a bargaining advantage from it. Rival record companies' inability to internalize the externality leads to weaker investement incentives, putting them at a disadvantage when bidding for an artist's new album. To the extent that other market frictions prevent the sale of an artist's entire catalog, the backward externality serves as a barrier to entry by locking artists into long-term relationships with their labels.

1 Introduction

The market for recorded music is famously concentrated: for the past decade, over 75% of the world's recorded music has been distributed by a handful of large, vertically integrated record companies.[1] Each of these companies owns several labels that operate more or less independently;

*We are grateful to seminar participants at Columbia, NYU-Stern, Penn State, Yale, Berkeley, University of Pennsylvania, and University of British Columbia for helpful comments and questions. We would especially like to thank Dirk Bergmann and Michael Whinston for their suggestions, which led to major changes in the paper. We are of course responsible for any errors and shortcomings that remain.

[1]The "big six" (Universal, Polygram, Sony, BMG, EMI, and Warner) became the "big five" when Universal acquired Polygram in 1998, and then the "big four" when regulators in the U.S. and E.U. approved the merger of Sony and BMG in 2004. McMillan [7] offers a brief narrative of the consolidation wave in his study of industrial clustering and the birth of reggae music.

for example, Warner Music Group has at least three major labels (Atlantic, Electra, and Warner) along with many smaller labels (such as Curb and Sire Records). In many cases these labels were originally independent, but when their artists starting producing major hits they were acquired by one of the major record companies.

Why is the recorded music market so concentrated? What prevents more efficient entrants from bidding for and winning the production rights to artists' new albums? One answer is that established artists are tied to their record labels by long-term contracts.[2] The difficulty with this explanation is that contract lengths are restricted (usually to seven years), and in any case these contracts do not appear to be binding: artists' contracts are almost always renegotiated after a hit album. Another possible explanation is that the vertically integrated distributors charge higher fees for distributing the entrant's albums, which gives them a competitive advantage over the entrant in bidding for new albums. However, the magnitude of the distribution fees, as well as the existence of some competition from alternative distribution companies, suggest this explanation is unlikely to be the entire story.

In this paper we offer another explanation based on an investment externality. The contracts under which artists produce albums are classified as "work-for-hire" agreements, so the labels own the copyrights on the albums. The release of a new album may significantly increase demand for the artist's earlier albums, known as the artist's catalog. We call this investment externality the *backward externality* to distinguish it from the *forward externality*, which refers to the impact of a new release on the demand for the artist's future albums. The label that owns the artist's catalog benefits from the backward externality. More importantly, it gives the label an incumbency advantage in bidding for the artist's new album: the incumbent's willingness to invest in the new album, and to pay for it, exceeds that of rival labels since the incumbent label internalizes the backward externality (whereas its rivals do not). Thus, even if artists and distributors cannot fully commit to long-term contracts, artists are nevertheless locked into long-term relationships with their labels. They will stay with a label even after it merges with or is acquired by a major record company. Entrants can compete for new albums by established artists only by purchasing the rights to their catalog, which is likely to be quite difficult since the negotiations involve rival labels rather than individual artists. The major recording companies will be reluctant to sell artists' catalogs if they believe that such sales would strengthen a rival against whom they have to compete in other

[2] Aghion and Bolton [1] have shown how private information can lead to outcomes in which an incumbent firm signs long-term contracts with a buyer to prevent entry by a more efficient firm.

markets. A potential entrant can only enter by developing its own stable of artists, which can take a long time and may be impossible to accomplish without a prior reputation from having produced hit artists.

Figure 1 illustrates the significance of the backward externality. The figure plots the logarithm of weekly national sales for the first and second albums of two popular recording artists, from the time of the artist's debut until six months after the artist's third release. The vertical lines in each graph indicate the release dates of the second and third albums. In the weeks surrounding these release dates, sales of catalog titles increase substantially. In the case of the "Bloodhound Gang," a relatively obscure alternative rock band, the second album was considerably more popular than the first, and its release catapulted sales of the prior album to levels even higher than it had attained at the time of its own release. For the "Foo Fighters," a more popular hard rock band with a very successful debut album, the impact of the second release is somewhat less dramatic, but still seems to have generated an increase in sales of the band's first album. In both examples, the backward externality is significantly positive. The effect appears to begin in the weeks just prior to the new album's release, and it persists for many months. In fact, for the "Bloodhound Gang," the effect persisted for at least three years.

Our main empirical goal in this paper is to measure the magnitude of the backward externality and show that it is large enough to affect the contractual relationships between artists and record labels. We measure the backward externality using a large sample of recording artists whose debut albums were released in the United States during the period 1993 to 2002. Many of the artists in our sample released as many as four albums during the sample period, allowing us to study the variation in backward externality over different pairs of albums (e.g., release of albums 3 and 4 on sales of albums 1 and 2, and album 4 on sales of album 3). A secondary goal of the paper is to identify the source of the investment externalities. Are the externalities due to preference learning or preference complementarities? The answer has important normative and policy implications.

Our empirical strategy for quantifying the backward externality is taken from the literature on treatment effects. The release of a new album is interpreted as "treatment," and we are interested in measuring the difference between catalog sales of the artist with treatment vs. without treatment. Since the artist cannot be in both states at the same time, we only observe one of the outcomes. However, if release times are random, then catalog sales for "untreated" artists with the same number of catalog albums can be used to estimate the counterfactual sales for "treated" artists. Of course, the variation in release times across artists may not be entirely random. We use fixed effects to control for time-invariant factors such as genre and artist popularity that may influence release times, and we also use a differences-in-differences estimator to control for possible correlation between the shape of the catalog album's sales path and the release time of the new album. We find that the average treatment effects are positive, permanent, and both statistically and economically significant.

We develop a simple model in which the investment externalities arise from consumer learning: the release of a new album increases the stock of potential consumers for the artist's future albums (the forward externality) and also generates new information about the quality of the artist's previous albums (the backward externality). The model predicts that the increase in catalog sales is always larger when the new release is a hit; the largest increases occur when a hit follows a non-hit and the smallest occurs when a hit is followed by a non-hit. The model also predicts that sales of a new release decline faster when it follows a hit because the stock of potential consumers is larger (i.e., the forward externality). We examine these predictions and find that they are consistent with the data. On the basis of these results and other facts about the timing of the treatment effects, we argue that the externalities are mainly due to preference learning rather than preference complementarities.

Although our model is specifically tailored to describe the music industry, we wish to emphasize that similar kinds of information spillovers are observed in various other markets. For example, an appearance on the bestseller list stimulates sales of an author's previous books, and a breakthrough performance by a film actor increases demand for other films in which the actor appeared. More generally, our model is related to the literature on brand extension, which analyzes firms' decisions about whether to release new products under existing brand names. (See, for example, Choi (1998), Cabral (2000), and Wernerfelt (1988).) When consumers are uncertain about product qualities, the strong reputation of an existing product increases demand for new products sold under the same brand (the forward externality), and the release of a high-quality new product can improve the brand image and boost sales of the existing product (the backward externality).[3] What distinguishes the music industry example is the aspect of joint production: the informational spillovers occur within the context of a bilateral contracting problem, with the externalities affecting the bargaining between artists and their record labels.

The paper is organized as follows. In Section 2 we present a model of investment externalities

[3] In Cabral's paper, for example, the "feedback reputation effect" is exactly analogous to what we call the backward externality.

and contracting. Section 3 discusses the data, and Section 4 outlines the empirical strategy for estimating the backward externality. Section 5 reports the estimation results. In Section 6 we test various predictions of the model and provide a quantitative assessment of the magnitudes of the investment externalities. Section 7 discusses alternative explanations of our findings. Section 8 concludes.

2 The Model

We develop a simple two period, two album model of investment externalities and contracting.[4] The main purpose of the model is to frame the analysis and interpret the results.

The standard contract between a new recording artist and a label is a combination of a debt and equity contract.[5] The artist agrees to deliver one album. In return, the label gives the artist a loan, called an advance, to finance living and studio costs. For the debut album, the allowance awarded typically ranges between $150,000 to $250,000. Occasionally, an artist has developed enough of a name for herself that she is able to generate competition among labels and demand a higher advance. The artist also receives a share of the album's sales revenues, known as royalties. The royalty rate for the debut album is between 10-14% of retail sales, with the typical number being 12%. The artist has to repay the advance out of her royalties. Thus, the label collects 100% of the album's revenues until its loan has been repaid, after which it shares the revenues with the artist based on the royalty rate. If revenues fall short of the advance, the label takes the loss. For an album to be successful, it needs to be marketed. The amount that a label often spends on the promotion and marketing of an album ranges between $200,000 to $500,000. These expenditures are not charged against the artist's royalties nor are they specified in the contract. Occasionally, the contract includes clauses about how or how much the artist will be promoted but they tend not to be very concrete and are difficult to enforce. In fact, the label may not even commit to releasing the first album. The contract is classified as a "work-for-hire" agreement so the label owns the copyright on the album.[6]

The label has the option to extend the term of the contract for more albums, typically five, at the same terms and conditions that apply to the first album. However, in practice, contract terms

[4] We are grateful to Michael Whinston for offering helpful insights in developing the model.
[5] See Krasilovsky et al's [6] book on contracts in the music industry.
[6] In some rare cases, the artists negotiate reversions, that is, ownership of the album reverts to the artist after some period of time.

are almost always renegotiated after the artist has a successful album, and both label and artist understand that from the outset. According to Don Engel, one of the more successful lawyers who specializes in renegotiating contracts, artists have a fair amount of bargaining leverage following a hit album. He claims that it is not difficult to get artists out of their contracts and has developed a number of legal strategies for making this happen. The main form of the renegotiation is a higher advance. Royalty rates may also get renegotiated up, but only by a few percentage points. The "gold" standard is 18% of retail. Artists whose albums did poorly just hope that the label will exercise its option for the next one. The artist has to repay any losses on the first album and the advance on the second album out of her share of sales revenues on *both* the first and second albums.

Given this description, we model the relationship between the artist and the label as follows. In period 1, Label I signs the artist to a contract and gives the artist an advance to cover living and studio costs, F. The artist produces an album of uncertain quality, which we model as a random draw z_1 from a marginal distribution G_{Z_1}. In period 1.5, Label I chooses its production and marketing investment i_1 and revenue $r(i_1, z_1)$ is realized. Since album sales are public information, z_1 is revealed to the market by the end of the period. In period 2, the artist auctions the rights to her second album, possibly switching to outside label E. After signing a new album contract with Label i, $i = I$ or E, nature gives the artist a random draw z_2 from the conditional distribution $G_{Z_2|Z_1}$. In period 2.5, Label i chooses its investment i_2 in album 2 and revenues $R_j(i_1, i_2; z_1, z_2)$ for $j = 1, 2$ are realized. The price of an album is p. It is the same for each album and fixed over time.

The random variables Z_1, Z_2 are assumed to be affiliated. The assumption implies that the family of conditional distributions $G_{Z_2|Z_1}$ are stochastically ordered in the first order sense and that the expected quality of album 2 is increasing in the quality of album 1. The contract for album 1 is assumed to be a royalty contract that gives the artist a share $\alpha \in (0, 1)$ of any revenues from album 1. The royalty rate on album 2 is assumed to be fixed, so negotiations are only over the fixed transfer. As mentioned above, the actual contract is somewhat more complicated since the label recoups its advance from the artist's royalties. We ignore this feature of the contract to simplify the presentation of the main ideas. The model is easily extended to incorporate the recovery of advances.

We can now define more precisely what we mean by investment externalities. Let $z = (z_1, z_2)$. Then, investment in album 2 generates a *backward externality* if $R_1(i_1, i_2, z)$ is strictly increasing in i_2. Similarly, investment in album 1 generates a *forward externality* if $R_2(i_1, i_2, z)$ is strictly increasing in i_1. Investments are complements (substitutes) if the revenue functions are supermodular (submodular) in (i_1, i_2).

The revenue functions depend upon demand. The market consists of a continuum of consumers. Each consumer's preferences are additive across albums by the same artists (and by different artists). Let u_{ij} denote consumer i's standalone utility of album j, $j = 1, 2$. It takes the form

$$u_{ij} = z_j \eta_{ij},$$

where η_{ij} is the idiosyncratic component of consumer i's utility. The η_{ij}'s are nonnegative, with mean one, correlated across albums, and independent of album qualities. Let H_j denote the marginal distribution of η_{ij}. Consumer i's purchase rule is to buy album j if and only if the utility of the album exceeds price. Hence, normalizing the number of consumers to 1, the (potential) demand for album j given album j quality is $1 - H_j(p/z_j)$. The additivity assumption implies that purchasing one album has no impact on the utility of a second album. It rules out demand externalities that can either lead to cannibalization or complementarity effects.

Consumers learn about their preferences for an album by hearing the album played on the radio or at concerts or at listening posts in music stores. We consider a very simple model of learning in which the time at which a consumer becomes aware of album 1 is random and independent of his preferences. The probability that a consumer learns about album 1 during period 1 is denoted by $q_1(z_1, i_1)$. It is an increasing function of album quality and marketing to reflect the fact that higher quality, more heavily promoted albums typically get more playing time. When a consumer learns about album 1, he knows his utility for the album. Hence, revenues for album 1 in period 1 is given by

$$r(z_1, i_1) = q_1(z_1, i_1)[1 - H_1(p/z_1)].$$

In period 2, demand for album 2 is determined by two kinds of consumers: informed consumers who have learned about the artist in period 1 and uninformed consumers who have not yet discovered the artist but may do so during period 2. Informed consumers are assumed to know about album 2 when it is released. The probability that an uninformed consumer learns about album 2 during period 2 is given by $q_2(z_2, i_2)$ denote. It is assumed to be an increasing function of album quality and investment. Then revenues for album 2 conditional on album qualities are given by

$$R_2(i_1, i_2, z) = (q_1(z_1, i_1) + q_2(z_2, i_2)(1 - q_1(z_1, i_1))[1 - H_2(p/z_2)].$$

When an uninformed consumer learns about album 2, then he also learns about album 1. Thus, revenues for album 1 in period 2 are given by

$$R_1(i_1, i_2, z) = q_2(z_2, i_2)(1 - q_1(z_1, i_1))[1 - H_1(p/z_1)].$$

The sources of the investment externalities in this model are informational. The backward externality occurs because some consumers do not learn about album 1 until period 2. The magnitude of the backward externality depends the capacity of album 2 to draw new consumers to the market, which in turn depends upon how well album 2 does and on the number of uninformed consumers. Uninformed consumers are more likely to discover the album during period 2 if album 2 is a hit than a dud. Hence, R_1 is increasing in z_2 and i_2. The stock of uninformed consumers is lower when album 1 is a hit, but the purchasing probability is higher. Thus, the effect of z_1 and i_1 on R_1 is ambiguous. The forward externality arises from consumers who discover the artist in period 1. The larger the stock of informed consumers, the higher is the demand for album 2. Hence, higher values of z_1 and i_1 increases R_2. Finally, R_2 is increasing in z_2 and i_2 since it raises the probability that uninformed consumers hear the album and, upon hearing it, purchase the album. Investments in our model are substitutes.

2.1 Contract Predictions

The backward externality explains a number of stylized facts about the contractual relationship between the artist and a label. If the artist switches to an outside label E in period 2, E's investment in album 2 will be

$$i_2^E(i_1, z_1) = \arg\max_{i_2} E_{Z_2|Z_1}[(1 - \alpha)R_2(i_1, i_2; z) - i_2].$$

If the artist stays with label I, investment will be

$$i_2^I(i_1, z_1) = \arg\max_{i_2} E_{Z_2|Z_1}[(1 - \alpha)(R_1(i_1, i_2; z) + R_2(i_1, i_2; z)) - i_2].$$

It is straightforward to show (see appendix) that i^I exceeds i^E if and only if R_1 is strictly increasing in i_2. The reason is that the incumbent label internalizes the impact of the release of the second album on album 1 sales, whereas the outside label does not. Hence, the incumbent label is always willing to invest more in album 2 than the outside label.

We show in the appendix that this result has several important implications. First, the outside label's willingness to pay for the rights to the second album is increasing in z_1, so it bids for the artist's second album only if the first album does well. This implication is consistent with the stylized fact that the contract renegotiations only take place following a hit album. If album 1 does poorly, the outside label is not willing to pay anything to acquire the rights to album 2, so the artist has no bargaining leverage. Second, even when the outside label is not willing to bid, the incumbent label is often willing to finance the second album because it internalizes the benefit of a second release on the revenues of the first album. This implication is consistent with the stylized fact that incumbents often release albums of artists who have not yet had a hit album. Third, when the outside label is willing to pay for the rights to the second album, the artist will always stay with the incumbent label. The outside label cannot bid away the artist because its investment incentives are not as strong as those of the incumbent label. This implication is consistent with the stylized fact that artists rarely switch labels. (In our sample, fewer than 10% of artists ever switch between major labels.) Note that the difficulty here is not a lack of commitment: even if the outside label could commit to matching the investment level of the incumbent label, the artist would not want the outside label to do so. The commitment would lower the outside label's willingness to pay and hence the amount that she can extract from the incumbent label.

The presence of a backward externality also explains the strategy adopted by the distributors when they decided to vertically integrating backwards into the production of music. Instead of bidding for the rights to new albums, they purchased the copyrights to catalog albums by buying labels. Ownership of an artist's catalog gives the new owners an incumbency advantage in "bidding" for the artist's new albums and ensures them that they will have a long-term relationship with the artists.

What is the impact of the investment externalities on investment in album 1? A larger investment in album 1 increases the stock of informed consumers available to buy album 2 and raises the outside label's willingness-to-pay for the rights to album 2. By threatening to switch to the outside label, the artist is able to capture these rents. This is the familiar holdup problem. If the artist is unable to commit not to renegotiate in the event of a hit, then the holdup problem reduces the willingness of a label to finance and invest in the debut album. The backward externality gives the incumbent rents that cannot be competed away by an outside label. An increase in these rents lowers the financing threshold for album 1, and it is in this sense that a positive backward externality mitigates the hold-up problem. But the impact of an increase in the magnitude of the backward externality on investment in album 1 depends upon the properties of the revenue functions. In our model, album investments are substitutes, so an increase in the backward externality lowers i_1 conditional on album 1 meeting the financing threshold.

2.2 Empirical Predictions

We do not directly observe the investment levels or album quality. This limits our ability to determine the properties of the reduced form revenue functions. However, we do observe when the second album is released, and the sales of album 1 before and after its release. Therefore, using sales as proxies for album quality, and treating i_j as a binary variable that is equal to one if the label invests in album j and zero otherwise, we can study the properties of

$$\Delta(z_1, z_2) = [R_1(1, 1; z_1, z_2) - R_1(1, 0; z_1, z_2)].$$

The key empirical issue is predicting the revenues of album 1 in the counterfactual world where album 2 is not released. Following the treatment literature, we will use album 1 sales of other, comparable artists to forecast $R_1(1, 0; z_1, z_2)$. If Δ is estimated to be positive, then the backward externality is present.

We also examine the variation in Δ with respect to z_1 and z_2. We partition albums into hits (H) and non-hits (N) and test the following two predictions of our model: (1) $\Delta(N, H) > \Delta(N, N)$ and (2) $\Delta(H, H) > \Delta(H, N)$.These predictions are quite robust and follow from the fact that the likelihood of uninformed consumers learning about the artist is higher when the second album is a hit. Holding constant the quality of album 2, the effect of increasing the quality of album 1 is ambiguous: hits lower the stock of uninformed buyers in period 2 but increase the probability that uninformed consumers buy album 1 when they learn about the artist. However, in our model, the first effect is likely to dominate. The reason is the forward externality: if album 1 is a hit, then album 2 is more likely to be a hit as well. Hence, the quality of album 2 is likely to be substantially higher if it is the first hit than if it is the second hit. Similarly, if album 2 is not a hit, then its quality is likely to be substantially lower if album 1 was a hit than if album 1 was not a hit. This argument suggests the following two predictions: (3) $\Delta(N,H) > \Delta(H,H)$ and (4) $\Delta(N,N) > \Delta(H,N)$.

The forward externality is not directly observable. However, it is possible to infer its presence and importance if one is willing to assume that the arrival rate of informed buyers is higher than that of uninformed buyers who have to learn about the artist during period 2. We test the following predictions: (5) the rate of decline in sales of album 2 is faster when album 1 is a hit and (6) the rate of decline in sales of album 2 is faster when album 2 is not a hit. Prediction (5) follows from the fact that the stock of informed consumers is higher when album 1 is a hit. Of course, if album 2 is not a hit, then the stock of uninformed consumers may not matter much since few will learn about the artist. In this case, the decline rate in album 2 will also be determined primarily by the arrival rate of informed buyers, which leads to prediction (6).

3 Data

Our data describe the album sales histories of 355 music artists who were active between 1993 and 2002. Weekly sales data for each artist's albums were obtained from Nielsen SoundScan, a market research firm that tracks music sales at the point of sale, essentially by monitoring the cash registers at over 14,000 retail outlets. SoundScan is the principal source of sales data for the industry, and is the basis for the ubiquitous Billboard charts that track artist popularity. Various online databases were also consulted for auxiliary information (e.g., about genres and record labels) and to verify album release dates.

The sample was constructed by first identifying a set of candidate artists who released debut albums between 1993 and 2002, which is the period for which SoundScan data were available. Sampling randomly from the universe of such artists is infeasible, largely because it is difficult to find information on artists who were unsuccessful. Instead, we constructed our sample by looking for new artists appearing on Billboard charts. The majority of artists in our sample appeared on Billboard's "Heatseekers" chart, which lists the sales ranking of the top 25 new or ascendant artists each week.[7] A smaller number of artists were found because they appeared on regional "New Artists" charts, and an even smaller number were identified as new artists whose debut albums went straight to the Top 200 chart. This selection is obviously nonrandom: an artist must have

[7] Artists on the Heatseekers chart are "new" in the sense that they have never before appeared in the overall top 100 of Billboard's weekly sales chart—i.e., only artists who have never passed that threshold are eligible to be listed as Heatseekers.

enjoyed at least some small measure of success to be included in the sample. However, although the sample includes some artists whose first appearance on the Heatseekers list was followed by a rise to stardom, we note (and show in detail below) that it also includes many unknown artists whose success was modest and/or fleeting. (The weekly sales of the lowest-ranked artist on the Heatseekers chart is typically around 3,000, which is only a fraction of typical weekly sales for releases by famous artists who have graduated from the Heatseekers category.)

Because our primary objective is to study demand responses to newly released albums, we restrict our attention to major studio releases. Singles, recordings of live performances, interviews, holiday albums, and anthologies or greatest hits albums are excluded from the analysis because they rarely contain any new music that could be expected to affect demand for previous albums.[8] The resulting sets of albums were compared against online sources of artist discographies to verify that we had sales data for each artist's complete album history; we dropped any artists for whom albums were missing or for whom the sales data were incomplete.[9] Since timing of releases is an important part of our analysis, we also dropped a small number of artists with albums for which we could not reliably ascertain a release date.[10] Finally, we narrowed the sample to artists for whom we observe the first 52 weeks of sales for at least the first two albums; we then include artists' third and fourth albums in the analysis if we observe at least the first 52 weeks of sales for those albums (i.e., we include third and fourth albums if they were released before 2002).

After applying all of these filters, the remaining sample contains 355 artists and 962 albums. The sample covers three broad genres of music: Rock (227 artists), Rap/R&B/Dance (79 artists), and Country/Blues (49 artists). The artists in the sample also cover a broad range of commercial success, from superstars to relative unknowns. Some of the most successful artists in the sample are Alanis Morissette, the Backstreet Boys, and Shania Twain; examples at the other extreme include Jupiter Coyote, The Weakerthans, and Melissa Ferrick.

[8] Greatest hits albums could certainly affect sales of previous albums—repackaging old music would likely cannibalize sales of earlier albums—but we are primarily interested in the impact of *new* music on sales of old music. Moreover, there are very few artists in our sample that actually released greatest hits albums during the sample period, making it difficult to estimate their impact with any statistical precision.

[9] The most common causes for missing data were that a single SoundScan report was missing (e.g., the one containing the first few weeks of sales for the album) or that we pulled data for the re-release of an album but failed to obtain sales for the original release.

[10] For most albums, the release date listed by SoundScan is clearly correct; however, for some albums the listed date is inconsistent with the sales pattern (e.g., a large amount of sales reported before the listed release date). In the latter case, we consulted alternative sources to verify the release date that appeared to be correct based on the sales numbers. Whenever we could not confidently determine the release date of an album, we dropped it along with all other albums by the same artist.

For each album in the sample, we observe weekly sales from the time of its release through the end of 2002. The key feature of the data is that sales are reported at the album level, so that we can observe the sales of prior albums when a new album is released. Both cross-sectional and time-series variation can be exploited to measure the sales responses: for a given album, we observe both that album's sales history prior to the new release and also sales paths for other comparable artists who did not release new albums.

Table 1 summarizes various important features of the data. The first panel shows the distribution of the albums' release dates separately by release number. The median debut date for artists in our sample is May 1996, with some releasing their first albums as early as 1993 and others as late as 2000. There are 74 artists in the sample for whom we observe 4 releases during the sample period, another 104 for whom we observe 3 releases, and 177 for whom we observe only 2 releases. Note that while we always observe at least two releases for each artist (due to the sample selection criteria), if we observe only two we do not know whether the artist's career died after the second release or if the third album was (or will be) released after the end of the sample period. In what follows we will discuss this right-truncation problem whenever it has a material impact on the analysis.

The second panel of the table illustrates the considerable heterogeneity in sales across albums. Production, marketing, and distribution costs for a typical album are in the ballpark of $500,000, so an album must sell roughly 50,000 units (assuming a wholesale price of $10 per unit) in order to be barely profitable; over half of the albums in our sample passed that threshold in the first year. However, although most of the albums in the sample were nominally successful, the distribution of success is highly skewed: as the table illustrates, sales of the most popular albums are orders of magnitude higher than sales of the least popular ones. For debut albums, for example, first-year sales at the 90^{th} percentile first release are ten times sales at the median and over 100 times sales of the album at the 10^{th} percentile.

The skewness of returns is even greater across artists than across albums, since artist popularity tends to be somewhat persistent. An artist whose debut album is a hit is likely to also have a highly successful second release, so absolute differences in popularity among a cohort of artists are amplified over the course of their careers. Across the artists in our sample, the simple correlation between first-year sales of first and second releases is 0.52. For second and third (third and fourth) releases the correlation is 0.77 (0.70). Most of an artist's popularity appears to derive from artist-specific factors rather than album-specific factors, but the heterogeneity in success across albums by a given artist can still be substantial.

Another interesting feature of the sales distributions is how little they differ by release number. To the extent that an artist's popularity grows over time, one might expect later albums to be increasingly successful commercially. However, while this pattern appears to hold on average for albums 1 through 3, even for artists who ultimately have very successful careers it is often the case that the most successful album was the first. In our sample, among the 74 artists for whom we observe four releases, 42 had the greatest success with either the first or second release.

Although there is obviously heterogeneity across albums, the "typical" sales path exhibits an early peak followed by a steady, roughly exponential decline. As indicated in the third and fourth panels of table, sales typically peak in the very first week and are heavily "front-loaded": a large fraction of the total sales occur in the first four weeks after release. Debut albums are an exception: first releases sometimes peak after several weeks, which presumably reflects a more gradual diffusion of information about albums by new artists. The degree to which sales are front-loaded seems to increase with each successive release.

Seasonal variation in demand for music CDs is substantial. Overall, sales are strongest from late spring through early fall, and there is a dramatic spike in sales during mid- to late-December. Not surprisingly, album release dates exhibit some seasonality as well. Table 2 lists the distribution of releases across months. Late spring through early fall is the most popular time to release a new album, and record companies appear to avoid releasing new albums in December or January. Albums that would have been released in late November or December are presumably expedited in order to capture the holiday sales period.

The last panel of table 1 summarizes the delay between album releases. The median elapsed time before the release of the second album is more than two years, and the low end of the distribution is still more than one year. The distribution of time between albums 2 and 3 is very similar. Fourth albums appear to be released more quickly, but this likely reflects sample selection. We can only compute time-to-next-release conditional on there being a next release, and since most of the third albums in our sample were released near the end of the sample period, we only observe a fourth release if the time to release was short. This right truncation applies to the other albums as well, but we do not expect the problem to be as severe in those cases. Figure 2 shows a more complete picture of the heterogeneity in release lags across albums, including elapsed time between non-adjacent albums. The distribution of elapsed time between albums 1 and 2 is clearly very similar to the distribution between albums 2 and 3, but the right truncation is obvious in cases involving the release of album 4.

In addition to the obvious right truncation problem, our sample selection is likely to be biased toward artists whose success came early in their careers. For an artist to be selected into our sample, it must be the case that (a) the artist appeared on a Billboard chart between 1993-2002, and (b) we have data on all the artist's CD sales, which means the artist's first release must have been after 01 Jan 1993. Taken together, these conditions imply that artists who hit a Billboard chart early in the sample period must have done so on their first or second album (otherwise we would have excluded them due to lack of data on their previous releases). Moreover, of the artists debuting late in our sample period, only the ones with early success will make it into our sample, because only they will have appeared on a Billboard chart. So the selection pushes toward artists who "start strong."

While this means our data will overstate the tendency of artists' successes to come early in their careers, we do not see any obvious biases the selection will induce in the empirical analyses of section 5. Moreover, a quick check of some out-of-sample data suggests the selection bias is not very severe. We compiled a list of 927 artists who appeared on the Heatseekers chart between 1997-2002 but who are not included in our sample. Of these artists, 73% made it to the chart on their first or second album, as compared to 87% for the artists in our sample. The difference is qualitatively consistent with the selection problem described above, but we do not think the difference is quantitatively large enough to undermine our main results.

4 Empirical Strategy

In this section we discuss our empirical strategy for estimating the backward externality. Our approach is taken from the treatment literature[11] and exploits exogenous variation in albums' release times.

A new album release by an artist is interpreted as the "treatment." Releasing a new album is an irreversible act: once treated, the catalog albums remain treated. We will follow the impact of a new release on sales of catalog albums for S periods, and refer to this number as the length of the treatment "window." (In the models estimated below, S is 39 weeks: 13 pre- and 26 post-treatment.) Each new release following the release of the debut album is analyzed as a separate

[11] See Wooldridge [10] for a summary.

treatment episode. For each episode, time is measured in terms of the number of periods since the last new album was released. The sample for each episode consists only of artists who receive treatment. Thus, the first episode includes the entire sample, since one of the criteria for selection is that the artist had to release at least two albums. The second and third episodes consists of artists who released three albums and four albums respectively.

Without loss of generality, we focus on the first treatment episode. Let y_{it}^0 denote the log of album 1 sales of artist i in period t without treatment, and let y_{it}^s denote the log of album 1 sales in period t when artist i is in the s^{th} period of treatment. Our objective is to estimate the average treatment effect on the treated (ATE) for each period of the treatment window. Our focus on the treatment effect on the treated mitigates the right truncation problems that are present in our sample. Notice that, by taking logs, we are implicitly assuming that treatment effects are proportional, not additive. There are two reasons for adopting this specification. One is that the distribution of album sales is highly skewed. The other is that the average treatment effect is likely to be nonlinear: a new release has a larger impact on total sales of catalog titles for more popular artists. By measuring the treatment effect in proportional terms, we capture some of this nonlinearity. However, it could bias our estimates of the treatment effects upwards since proportionate effects are likely to be higher for less popular artists, and there are many more of them. Proportionate effects may also be higher for popular artists who are treated later since their sales levels are likely to be a lot lower than popular artists who are treated earlier. We address these issues in discussing the results below.

The main challenge in estimating the ATE's for artist i is that, in each period, we observe only one outcome for that artist. The observed outcome for artist i in period t is

$$y_{it} = y_{it}^0 + \sum_{s=1}^{S} w_{i,t-s+1}[y_{it}^s - y_{it}^0],$$

where $w_{i,t-s+1}$ is an indicator variable that is equal to one if artist i enters treatment in period $t-s+1$ and zero otherwise. The probability model generating outcomes for artist i in period t is given by:

$$y_{it}^s = \mu^s + \phi(t) + \nu_i + v_{it}^s,\ s = 0, 1, 2, .., S.$$

Here μ^s is the mean of the distribution of log sales in time period t for artists in the s^{th} period of treatment, $\phi(t)$ is a function that that captures the common, downward trend in an artist sales,

ν_i measures the impact of unobserved artist characteristics on sales in every period, and v_{it}^s is the idiosyncratic shock to album 1 sales of artist i when she is in treatment period s at time period t. The artist-specific effect does not vary across the treatment window. Substituting the above equations, the observed outcome for artist i in period t is given by

$$y_{it} = \mu^0 + \phi(t) + \nu_i + v_{it}^0 + \sum_{s=1}^{S} w_{i,t-s+1}[(\mu^s - \mu^0)] + (v_{it}^s - v_{it}^0)].$$

The ATE for treatment period s is the difference in means, $\mu^s - \mu^0$.

Intuitively, our strategy for measuring this difference is to use the sales of not-yet-treated albums (i.e., albums whose artists have not yet released a newer album) as the benchmark against which to compare sales of treated albums (i.e., albums whose artists have recently released new albums). Our specific sampling and estimation procedure is as follows. For each artist, t indexes time since the debut album's release, not calendar time. Albums are included in the sample only until the last period of the treatment window: observations on sales *after* that window are not used in estimating the regressions. We adopt this approach to ensure that, at any given t, treated albums are being compared with not-yet-treated albums, rather than a mix of not-yet-treated and previously-treated albums. Thus, the sample in period t includes artists that have not yet released a new album and artists who had a new release in periods $t-1, t-2, .., $ or $t-S+1$ but excludes artists whose new release occurred prior to period $t-S+1$. Basically, we want the control group to measure what happens to sales over time before any new albums are released: our approach assumes that for an album whose artist issues a new release at t, counterfactual sales (i.e., what sales would have been in the absence of the new release) can be inferred from the sales of all other albums at t for which there has not yet been a new release.[12]

The regression model is as follows:

$$y_{it} = \alpha_0 + \alpha_i + \lambda_t + \sum_{m=2}^{12} \delta_m D_{it}^m + \sum_{s=-13}^{25} \beta_s I_{it}^s + \epsilon_{it}, \tag{1}$$

where α_i is an artist fixed effect, the λ_t's are time dummies, and the D^m's are month-of-year dummies (to control for seasonality). Here I_{it}^s is an indicator equal to one if the release of artist i's new album was s weeks away from period t, so β_s measures the new album's sales impact in

[12]We believe dropping post-treatment observations is the most appropriate approach, but it turns out not to matter very much: our estimates change very little if we include these observations.

week s of the treatment window. ($t = 0$ corresponds to the first week following the new release.) Intuitively, after accounting for time and artist fixed effects, we compute the difference in the average sales of album 1 between artists in treatment period s and artists who are not treated for each period, and then average these differences across the time periods. The stochastic error, ϵ_{it}, is assumed to be heteroskedastic across i (some artists' sales are more volatile than others') and autocorrelated within i (random shocks to an artist's sales are persistent over time).

The time dummies (λ_t) allow for a flexible decay path of sales, but implicitly we are assuming that the shape of this decay path is the same across albums: although differences in the level of demand are absorbed in the album fixed effects, differences in the shapes of albums' sales paths are necessarily part of the error (ϵ). Including separate indicators for successive weeks of treatment allows us to check whether the new release's impact diminishes (or even reverses) over time, which is important for determining whether the effects reflect intertemporal demand shifts. We allow for a 39-week treatment window, beginning 13 weeks (3 months) *before* the release of the new album. The pre-release periods are included for two reasons. First, much of the promotional activity surrounding the release of a new album occurs in the weeks leading up to the release, and we want to allow for the possibility that the backward externality reflects consumers' responses to these pre-release marketing campaigns. Second, including pre-release dummies serves as a reality check: we consider it rather implausible that a new album could have an impact on prior albums' sales many months in advance of its actual release, so if the estimated "effects" of the pre-release dummies are statistical zeros for months far enough back, we can interpret this as an indirect validation of our empirical model.

The critical assumption for the regression described above to yield consistent estimates of the treatment effect is that the treatment indicators in a period are independent of the idiosyncratic sales shocks in that period. In other words, after controlling for time-invariant characteristics such as genre and artist quality that affect the level of sales in each period, we need the treatment to be random across artists. This is a strong but not implausible assumption. We suspect that the main factor determining the time between releases is the creative process, which is arguably exogenous to time-varying factors. Developing new music requires ideas, coordination, and effort, all of which are subject to the vagaries of the artist's moods and incentives. To better understand the sources of variation in release times, we estimate Cox proportional hazard models with various album and artist characteristics included as covariates. Table 3 presents the results. Somewhat surprisingly, the time it takes to release an artist's new album is essentially independent of the success of the prior album (as measured by first six months' sales) or its decline rate after conditioning on genre. Release lags are significantly shorter for Country artists, and the coefficients on "years since 1993" reveal a general time trend toward longer lags between second and third (and third and fourth) albums.

Nevertheless, the specific question for our analysis is whether release times depend on the sales patterns of previous albums in ways that album fixed effects cannot control. One possibility is that release times are related to the *shape* of the previous album's sales path. Although the insignificant coefficients on the decline rate variable in Table 3 seem to suggest that release times are unrelated to decline rates, subtle relationships between sales-path shapes and release times may still exist. For example, albums of artists that spend relatively more effort promoting the current album in live tours and other engagements will tend to have "longer legs" (i.e., slower decline rates) and later release times than albums of artists that spend more time working on the new album. It is also possible that release times vary for strategic reasons. If the current release is not a hit, record companies may delay investing in a new release until more information becomes available. In some cases artists may delay the production of new music as a bargaining tactic. (Most recording contracts grant the record company an option to produce future albums by the artist under the same terms as applied to previous albums. Artists' leverage for negotiating more favorable terms in these contracts derives partly from a threat to withhold new music.) Whatever the reason for the relationship between the shape of the sales path and the time to the next release, the potential problem is that our regression only controls for the average rate of decline in album sales, so our estimates of the treatment effect will be biased if deviations from that average are systematically related to release times.

In order to address this issue, we estimate the regression model of equation (1) using the first difference of *ln(sales)* as the dependent variable: i.e., we estimate

$$\Delta y_{it} = \widetilde{\alpha}_0 + \widetilde{\alpha}_i + \widetilde{\lambda}_t + \sum_{m=2}^{12} \widetilde{\delta}_m D_{it}^m + \sum_{s=-13}^{25} \widetilde{\beta}_s I_{it}^s + \widetilde{\epsilon}_{it} \ , \tag{2}$$

where $\Delta y_{it} \equiv y_{it} - y_{it-1}$. This model estimates the impact of new releases on the percentage rate of *change* (from week to week) in previous albums' sales. The advantage of this specification is that heterogeneity in sales levels is still accounted for (the first differencing sweeps it out), and the fixed effects, $\widetilde{\alpha}_i$, now control for unobserved heterogeneity in albums' decline rates. Taking this heterogeneity out of the error term mitigates concerns about the endogeneity of treatment with respect to the shape of an album's sales path.

5 Results

We estimate the regression in (1) separately for each of six "treatments": the impact of the second, third, and fourth releases on sales of the first album; the impact of the third and fourth releases on sales of the second album; and the impact of the fourth release on sales of the third album. In constructing the samples for estimating the regression we impose several restrictions. First, we exclude the first eight months of albums' sales histories, in order to avoid having to model heterogeneity in early time paths. Recall that although most albums peak very early and then decline monotonically, for some "sleeper" albums we do observe accelerating sales over the first few months. By starting our sample at eight months, we ensure that the vast majority of albums have already reached their sales peaks, so that the λ_t's have a better chance at controlling for the decay dynamics. For later treatments, we restrict the sample to begin eight months after the release of the previous album. So, for example, in estimating the impact of album 4 on album 2, we use album 2's sales beginning eight months after the release of album 3. In essence, we want to consider the impacts of the various releases separately, in each case taking the flow of sales just prior to the new release as given. A second restriction involves truncating the other end of the sales histories: we exclude sales occurring more than four years beyond the relevant starting point. This means that if an artist's second album was released more than four years after the first, then that artist is not included in the estimation of the impact of second releases on first albums, and (similarly) if an artist's third release came more than four years after the second, then that artist is excluded from the two regressions estimating the impact of album 3 on albums 1 and 2.

Table 4 presents estimates of equation (1), with standard errors corrected for heteroskedasticity across artists and serial correlation within artists. (Estimated AR(1) coefficients are listed at the bottom of the table.) The columns of the table represent different treatment episodes (album pairs), and the rows of the table list the estimated effects for the 39 weeks of the treatment window (i.e., the $\hat{\beta}_s$'s). Since the dependent variable is the logarithm of sales, the coefficients can be interpreted as approximate percentage changes in sales resulting from the new release.

In each treatment episode, the estimated impact of the new album three months prior to its actual release is statistically indistinguishable from zero. As discussed above, this provides some reassurance about the model's assumptions: three months prior to the treatment, the sales of soonto-be-treated albums are statistically indistinguishable from control albums (after conditioning on album fixed effects and seasonal effects). In general, small (but statistically significant) increases start showing up 4-8 weeks prior to the new album's release, growing in magnitude until the week of the release ($t = 0$ in the table), at which point there is a substantial spike upward in sales.

Figure 3 shows the estimated effects graphically: each panel plots the estimated coefficients (and 95% confidence bands) for one of the album pairs. As can be seen in the figure, the estimates of the effects for each of the weeks following the release of a new album are always positive, substantive, and statistically significant. The largest externality is between albums 2 and 1, with estimates ranging between 40-55%. The externalities for the remaining pairs of albums are smaller, ranging mostly between 15-35%; however, overall the magnitudes are remarkably similar between album pairs.

For most album pairs (and especially the impact of album 2 on album 1) the estimated effects are remarkably persistent: the externalities do not appear to be transitory. The only apparent exception is the impact of album 3 on album 2, for which the coefficients decline somewhat at the end of the treatment window. It is important to note, however, that the increasing coefficients in some specifications do not imply ever-increasing sales paths, since the treatment effects in general do not dominate the underlying decay trend in sales. (In order to save space, the table does not list the estimated time dummies, which reveal a steady and almost perfectly monotonic decline over time.)

Table 5 lists estimated coefficients from the first-differenced regression (2) for the six album pairs considered in Table 4, and Figure 4 shows the estimated patterns graphically. (The figure plots the cumulative impact implied by the estimated weekly coefficients. For comparison, the dashed line indicates the estimated effects from the levels regression (1).) The implied effects are qualitatively and quantitatively very similar to those reported in Table 4. Given this reassuring comparison, we are confident that our main results are driven by real effects, not by subtle correlations between current sales flows and the timing of new releases.

We also checked the robustness of the estimates by splitting the sample in each treatment based on the median treatment time. As expected, the patterns are the same but the estimated effects are smaller for the albums that are treated early and larger for albums treated later.[13] The estimates are

[13]This pattern makes sense because our model assumes the effects are proportional: albums treated later will tend to have lower sales flows at the time of treatment, so the proportional impact of the new release will tend to be larger than for albums with high sales flows.

always strongly significant.

6 Magnitudes and Predictions

The estimates described above indicate that new releases lead to large percentage increases in sales of catalog titles by the same artist, and the increases appear to be persistent. However, the absolute number of additional sales may be small if the percentage increase applies to trivially low sales flows. For example, the median time of release for the second album is at 107 weeks from the first release, and the median sales flow fourteen weeks prior (i.e., just before the treatment period in our empirical model) is 307, with a median decline rate of 1 percent per week. If we apply our estimated percentage increases to the implied sales flows for this hypothetical median album, the predicted total increase (over the 39-week treatment period) is only 3,012 units. Multiplied by a price of $16 per unit, this implies an externality of roughly $48,200—not a trivial number, but perhaps not large enough to have a substantial impact on contracting.

However, our model suggests important asymmetries in the magnitudes of the externalities. Recall from Section 2 that the backward externality is predicted to be large when the new release draws lots of new fans to the artist, which is likely to be the case if it is a hit and the stock of uninformed consumers is low; it is predicted to be negligible if the new release is a dud and/or most consumers already know about the catalog albums. To examine the variation in the backward externality, we split our sample depending on whether albums were "hits." We define a hit as an album that sold 250,000 units or more in its first year; 30% of the albums in our sample meet this criterion.[14] We then divide our sample into four categories—hits followed by hits, hits followed by non-hits, non-hits followed by hits, and non-hits followed by non-hits—and summarize the backward externalities for each of the four categories in Table 6.

The table is based on estimates of the regression model computed separately for each subgroup.[15] These are then used to calculate the implied total change in sales for the "median" album. Specifically, we calculate the median weekly sales 14 weeks prior to the median release time, and the median weekly decline over the 39 weeks that follow. (In these calculations, we use only albums

[14] As a point of reference, the RIAA certifies albums as "Gold" if they sell more than 500,000 units. Also, among the albums we categorize as hits, at least 90% had peak sales high enough to appear on Billboard's Top 200 chart (vs. less than 10% among those we categorize as non-hits).

[15] We use the first-differences model in equation (2). Some of the estimated sales increases are smaller if we estimate the model in levels, but the qualitative patterns are essentially the same.

whose artists have not yet released the next album, so that the median sales flows and median decline rates will not reflect any of the backward externalities.) For example, in the group of 53 artists whose first two albums were both hits, the median time between the first and second releases is 108 weeks. Among first albums for which there was not yet a second release, the median weekly sales at week 94 (=108-14) was 1,888, and the median decline rate over weeks 95-134 was 1.8% per week. So we take a hypothetical album, with weekly sales beginning at 1,888 and declining at 1.8% per week, and apply the percentage increases implied by our estimated coefficients. The predicted total increase in sales over the 39-week period is 23,766, or roughly $380,000 in additional revenues (again using a rough price of $16 per unit).

The patterns in Table 6 are consistent with the predictions of the model. The backward externality is always larger when the new album is a hit, whether the debut album was a non-hit (prediction (1)) or a hit (prediction (2)). The largest increase occurs when a non-hit album is followed by a hit (prediction (3)). Presumably artists in this category generated a larger stock of uninformed consumers, and a relatively higher quality second album, than artists with two successive hits. Note that sales flows at the time of the new release are much lower for the former group than for the latter, so the larger total increase reflects dramatically larger percentage increases. For an artist whose second album was her first hit, we estimate that weekly sales of her first album roughly double when the new album is released.[16] Finally, the smallest increase occurs when a hit is followed by a non-hit (prediction (4)).

The same patterns hold when we examine the impact of the third release on the sales of album 2. The externalities are large when the new album is a hit, but negligible otherwise. The numbers are slightly smaller than those for the previous album, which could be interpreted to reflect a shrinking stock of uninformed consumers. (By the time a third album is released, a larger fraction of an artist's potential market has become aware of or familiar with the artist's music.)

An important lesson from Table 6 is that although on average (across all types of albums) the backward externalities may seem economically unimportant, they are in fact quite large for the artists that matter: the ones that have hits or have the potential to produce hits. If the artist's next release has the potential to be a hit, then the backward externality will have a meaningful impact on the contracting relationship between artist and record label. On the other hand, if it is clear that the artist's career has peaked, or that most of the artist's potential market is already aware of the

[16]The median decline path is essentially flat in this group, which also contributes to the prediction of a large total sales increase. However, even if we impose a 2% weekly decline, the predicted total increase is still over 13,000 units.

artist, then no externality will be expected, and consequently there is no lock-in effect.

Table 7 provides evidence on the forward externality by reporting the ratio of sales in the first month to sales in the first year of a new release for each of the four categories. The ratio is a measure of how front-loaded an album's sales are. In Section 2 we argued that sales of new albums should be considerably more front-loaded when the previous album was a hit. If an artist's previous album was a hit, then there is a large stock of consumers in the market who are aware of the artist and his music. Assuming these consumers make their purchase decisions regarding the new album more quickly than uninformed consumers (e.g., because they know about the release date and visit the music store sooner thereafter), they will cause sales of the new album to be more front-loaded; so the more successful the previous album, the more front-loaded will be the sales of the new album (prediction (5)). This is exactly the pattern shown in Table 7. For both second and third releases, sales are significantly more front-loaded when the previous album was a hit. Also, sales are the least front-loaded for hit albums that were preceded by non-hits, which suggests success may diffuse more slowly when most consumers were previously unaware of the artist (prediction (6)).

The inference that externalities result from consumer learning has important implications. Interpreted literally, our model of consumer learning says that many artists sell far fewer albums than they would if consumers were fully informed. Note that, in our model, preferences are strictly additive across albums by the same and other artists, so there is no meaningful substitution between albums. Hence, artists sell fewer albums than they would have in the but-for world of complete information, but they do not lose sales to other artists. By comparing the ratios reported at the bottom of each panel of Table 6, we can get a crude estimate of the "lost" sales resulting from uninformed consumers. The table reports the average ratio of [cumulative sales prior to a new release] over [cumulative sales over four years]. So, for example, for artists whose first two albums were both hits, on average 73% of the first album's four-year sales come before the second album is released. In contrast, when the second album was not a hit—i.e., where we expect no backward spillovers—this ratio is 84%. If we attribute the difference entirely to the spillovers described in our model, a back-of-the-envelope calculation indicates albums that do not benefit from the backward externality "lose" about 13% of their total potential sales.[17] In other words, if the new release does not generate any informational spillovers, then the catalog album will sell 13% fewer total units than it

[17] Let M denote cumulative sales at the time of the new release, and let T_1 and T_0 denote total sales (over four years) of an album that benefits or does not benefit, respectively, from the backward externality. Then $M/T_1 = 0.73$ and $M/T_0 = 0.84$, so $T_0/T_1 = .73/.84 = .869$.

could have if the additional consumers had been informed. While crude, this calculation is roughly comparable to the estimates yielded by the (more careful) empirical analysis above.

7 Discussion

In our theoretical model and in our interpretation of the empirical results, the source of the externalities is assumed to be informational. Consumers have to learn about their preferences for albums either because they are not aware of the artist or too uncertain about whether they will like the album enough to buy it. The release of a new album generates new signals (e.g., new tracks are played on the radio), so new releases enhance awareness and reduce uncertainty. Albums that lead to favorable updates of beliefs about past albums will generate a backward externality. Are there other explanations for why sales of an artist's catalog albums increase when the artist releases a new title, especially when the new release is a hit?

7.1 Price Effects

One possible explanation is prices: if retailers routinely lower the prices of catalog albums by artists with new releases, then the increase in sales that we have been calling the backward externality is in fact not an externality; it simply reflects downward-sloping demand curves. We do not have any publicly available price data for the albums in our sample, so we cannot directly rule out this explanation; however, we have various reasons to doubt that the backward externalities are driven by discounts. First, retail CD prices are remarkably rigid in general. Variation in price across titles and over time is limited, with almost all albums being simply classified into a few pricing tiers. Discounts are occasionally "pushed down" to the retail level by distributors, but these discounts are usually for new albums rather than catalog titles. According to two retail store managers with whom we had conversations, even when catalog albums are discounted, the timing of the sales does not seem to be systematically related to new releases by the same artist.

Ideally we'd like to know about discounting practices during our sample period, but we do not have meaningful data on prices from that period.[18] However, one thing we can check is whether current retailers engage in discounting practices that could generate backward externalities. To this end,

[18] The Federal Trade Commission has collected price data for our sample period in order to study the proposed merger between Time-Warner and EMI in 2000. We are currently working to have those data released.

we identified 20 artists who released new titles in February 2005, and downloaded Amazon.com's prices on each of these artist's bestselling CDs. (We excluded singles, imports, and EP's, as well as a few high-priced special-edition anthologies.) To assemble a control group, we matched each of the 20 artists to two "similar artists" (where the similarity was suggested by Amazon.com), and again obtained prices for the bestselling CDs of each of these artists. This yielded 20 artists with recent releases, 40 artists without recent releases, and a total of 264 albums combined.

Table 8 compares the prices for three groups of albums: new releases, catalog titles by artists with new releases, and catalog titles by artists without new releases. The table clearly indicates that although new releases tend to be discounted, the price distributions for the other two groups are quite similar. A Kolmogorov-Smirnov test fails to reject that the prices of catalog titles have the same distribution regardless of whether the artist has a new release, and a simple t-test fails to reject that the mean prices are equal. Another interesting feature of Amazon.com is that two prices are typically posted: the "list price" and the Amazon.com price. If we define discounting to mean offering a price below list price, we find that 91 of 160 (57%) catalog titles without new releases were discounted, as compared to 47 of 84 (56%) catalog titles by artists who did have new releases. These numbers are obviously based on a small sample, and they describe pricing patterns over three years after the end of our sample period. However, we consider it highly unlikely that retail pricing policies in the 1990's were somehow more fluid or sophisticated than in 2005, and we are confident that the backward externalities estimated in section 5 are not merely responses to discounting.

7.2 Preference Complementarities

A second explanation that we consider more plausible is preference complementarities. Suppose consumers know their preferences for an artist's albums, but their utility functions are supermodular. The externalities would then result from direct complementarities in the consumption of albums by the same artist: i.e., owning one album by an artist endogenously increases the utility from purchasing other albums by that artist. This is essentially a story about "fans": when consumers listen regularly to an artist's music, they become accustomed to it or invested in the image associated with it, and therefore more likely to purchase more music from that artist. In this framework, the forward externality arises because album 1 increases the incremental utility of consuming album 2, and the backward externality occurs because album 2 raises the incremental utility of consuming album 1. Hence, releasing album 2 makes album 1 more attractive to consumers whose utility from both albums is high enough to induce purchase, but whose utility from album 1 alone was below the purchase threshold. In a previous version of this paper we developed a two album, two period choice model with preference complementarities which generated revenue functions with essentially the same qualitative properties as those described in Section 2.

Like the learning model outlined in Section 2, the preference complementarity model predicts persistent backward externalities, since in both cases the new release directly changes the probability of wanting to purchase the catalog album. Other information-based explanations may predict transitory demand increases, as the “buzz” and increased airplay around the time of a new release could accelerate the arrival of consumers at the store. However, if these consumers are ones who would have eventually purchased the catalog title anyway (i.e., even if the new album were never released), then the increases could not be as persistent as the ones we see in our data. In order to generate permanent increases in demand for past albums, a new release must induce purchases by customers who would not have otherwise purchased.

While the preference complementarity model is consistent with the persistence we observe in the sales increases, it has difficulty explaining several other features of the data. One is the presence of pre-release effects. Record companies commonly concentrate their marketing and promotion efforts in the weeks leading up to a new release, and often release singles to be played on the radio several weeks before the release of the full album. In the preference learning model, pre-release promotion naturally leads some consumers to update their beliefs about catalog albums, inducing some to go out and buy the old album even before the new one is released. Conversely, the preference complementarity model, strictly interpreted, suggests that the benefits of joint consumption can’t be obtained until both albums are available.

A second feature is the variation in decline rates across the four categories. In the preference complementarity model, one would have to argue that, for any new release, the arrival rate of “old” buyers (i.e. consumers who had bought at least one of the artist’s previous albums) is greater than the arrival rate of “new” buyers (i.e. consumers who have not yet bought any of the artist’s albums). However, it is not clear why this distinction between consumers should matter in a choice model with full information, where consumers know about the albums that have been released and know about their preferences for those albums.

The third feature relates to the timing of artists’ hits. In the preference complementarity model, predictions about the magnitudes of the backward externality depend on the distribution of preferences in the population. Sales of any single album are a relatively small fraction of the number of potential consumers, even for big hits. So preferences for an album must be highly skewed, with purchases coming from consumers in the right tail of the distribution. Under this assumption, the backward externality is much larger when the catalog album and new release are both hits than when they are both duds, because there is more mass on the marginal types for hits than for duds. In the latter case, the marginal types are in the tails of the joint distribution where there is very little probablity mass, particularly when the consumer's preferences are correlated across albums. The asymmetric cases are also easily ranked. If album 2 is a hit, then lots of consumers who did not buy album 1 in period 1 will be willing to buy it together with album 2 due to the complementarity effect. If album 2 is a dud, then few consumers buy album 2 and, as a result, even fewer change their minds about the utility of album 1. The more difficult comparison is two hits versus a dud followed by a hit. The ranking depends upon functional forms, but our analysis suggests that the largest externalities are likely to occur when the catalog title and new release are both big hits.

Overall, although we cannot claim the data provide a definitive answer, we believe the facts favor the preference learning model over the preference complementarity model. This distinction is not especially relevant to the central point of our paper: the impact of the backward externality on contracting and market structure is the same regardless of the externality's source. However, we think the distinction is still important, because if the consumer learning model is correct then it can help explain various other features of the market for recorded music, such as the dramatic skewness in the distribution of returns across artists and albums.

8 Conclusion

We find that the backward externality is on average positive, substantial, and permanent. The release of an artist's second album increases sales of the artist's debut album by approximately 40-50% per week, and the effects show little indication of declining even after six months. The impacts of the third and fourth releases on catalog albums are smaller (approximately 20-30% per week) but still significant. The estimated percentage effects imply total increases that are economically meaningful, particularly on the most recent catalog album.

The incidence of the backward externality is strongly related to the relative success of the catalog and new release. Indeed, the backward externality appears to be largely a hit-driven phenomenon. When album 2 is the first hit, album 1 sales begin increasing two months prior to album 2's release, and the percentage increases in the post-release period range from 60–100%. By contrast, when album 1 is a hit, the impact of album 2 on album 1 sales is relatively weak. Pre-release sales increases are negligible, and post-release increases range from 6-15% per week. The impact of the third album on the second is similar. These results suggest that the source of the backward externality is preference learning rather than preference complementarity.

The backward externality has important policy implications. The Recording Industry Association of America (RIAA) and American Federation of Television and Radio Artists (AFTRA) have repeatedly lobbied Congress to end long-term contracting, as was done in the movie industry in the 1940s (see Terviö (2004)). Our results suggest that eliminating the label's option to extend the terms of the contract for more albums might have little effect, since the lock-in effect associated with the backward externality could offset the hold-up problem associated with the forward externality.

Another important implication of our results is that the backward externality may present a barrier to entry. Entrant labels cannot internalize the externality, so they are at a disadvantage whenever the externality is likely to be important. Our results suggest that an entrant can successfully bid for new releases of artists whose careers are on the decline, but not for new releases by artists whose careers are on the rise. This is not a good situation for the entrant, particularly if the incumbent label is better able to forecast the artist's peak. In principle, the entrant could try to internalize the externality by purchasing the rights to the artist's catalog from the incumbent label; however, information asymmetries and strategic concerns will tend to prevent such trades. Because the incumbent label has private information about the artist, the usual adverse selection problems will inhibit the trading of artist's catalogs. From a strategic perspective, incumbent labels are unlikely to sell an artist's catalog if doing so facilitates the entry of a firm that will become its competitor in the market for music and in the market for new artists.

Finally, our empirical results suggest that some of the skewness in the distribution of sales across albums and artists is due to the process through which consumers learn about their preferences. Consumers learn about albums by listening to them on radio and at concerts, and by talking to friends about albums they have heard and concerts they have attended. Consequently, if consumers only buy what they hear, and they only hear what others buy, then albums that fail commercially are typically undersold (in the sense that many would-be buyers remained uninformed), and hits may be oversold. We intend to explore this issue more fully in a subsequent paper.

References

[1] Aghion, P. and P. Bolton, “Contracts as a Barrier to Entry,” *The American Economic Review*, 77, 3, June, 1987, pp. 388-401.

[2] Cabral, L.M.B., “Stretching Firm and Brand Reputation,” *Rand Journal of Economics,* 31, 4, Winter 2000, pp. 658-673.

[3] Choi, J.P., “Brand Extension as Informational Leverage,” *Review of Economic Studies,* v. 65 (1998), pp. 655-669.

[4] Gandal, N., Sarit Markovich, Michael Riordan, “Ain’t it ‘Suite’: Strategic Bundling in the PC Office Software Market” , mimeo, April, 2004.

[5] Gentzkow, M., 2004, “Valuing New Goods in a Model with Complementarity: Online Newspapers”, mimeo, 2004.

[6] Krasilovsky, M.W., Sidney Shemel, John M. Gross, 2000, *This Business of Music: The Definitive Guide to the Music Industry,* 9th Edition, Billboard Books.

[7] McMillan, J., “Trench Town Rock: The Creation of Jamaica’s Music Industry,” mimeo, June, 2005.

[8] Terviö, M., “Mediocrity in Talent Markets,” University of Berkeley working paper, December, 2004.

[9] Wernerfelt, B., “Umbrella Branding as a Signal of New Product Quality: An Example of Signalling by Posting a Bond.” *RAND Journal of Economics*, v. 19 (1988), pp. 458-466.

[10] Wooldridge, J., 2002, *Econometric Analysis of Cross-section and Panel Data*, MIT Press.

Appendix

In this appendix we establish the claims made in Section 2.

Claim 1: $i^I(i_1, z_1) > i^E(i_1, z_1)$ *if and only if* R_1 *is strictly increasing in* i_2*.*

Proof. Conditional on paying the fixed fee to finance the production of album 2, Firm E's optimal investment solves the following equation:

$$(1-\alpha)E_{Z_2|Z_1}\left\{\frac{\partial R_2(i_1, i_2^E, z)}{\partial i_2}\right\} - 1 = 0.$$

Similarly, Firm I's optimal investment solves the first order conditions

$$(1-\alpha)E_{Z_2|Z_1}\left\{\frac{\partial R_1(i_1, i_2^I, z)}{\partial i_2} + \frac{\partial R_2(i_1, i_2^I, z)}{\partial i_2}\right\} - 1 = 0.$$

Since revenues of album 2 is increasing in i_2, the result follows if and only if album 1 revenues are increasing in i_2. Q.E.D. ■

Claim 2: z_1^E *is unique.*

Proof. Define Label E's net profits from album 2 as

$$\pi^E(i_1, z_1) = E_{Z_2|Z_1}(1-\alpha)R_2(i_1, i_2^E, z)\} - i_2^E.$$

Here z_E solves

$$\pi^E(i_1, z_1^E) - F = 0.$$

We need to show that Label E's net profits are increasing in z$_1$. Differentiating π^E with respect to z_1 yields

$$\begin{aligned}
\frac{\partial \pi^E}{\partial z_1} &= (1-\alpha)E_{Z_2|Z_1}\left\{\frac{\partial R_2}{\partial z_1} + \frac{\partial R_2(i_1, i_2^E, z)}{\partial i_2}\frac{\partial i_2^E}{\partial z_1} + \frac{\partial E_{Z_2|Z_1}\left\{R_2(i_1, i_2^E, z)\right\}}{\partial z_1}\right\} - \frac{\partial i_2^E}{\partial z_1} \\
&= (1-\alpha)\left[E_{Z_2|Z_1}\left\{\frac{\partial R_2}{\partial z_1}\right\} + \frac{\partial E_{Z_2|Z_1}\left\{R_2(i_1, i_2^E, z)\right\}}{\partial z_1}\right].
\end{aligned}$$

The second equality follows from the envelope theorem. The forward externality implies that the first term in the brackets is positive. The second term is positive because R_2 is increasing in z_2 and Z_1 and Z_2 are affiliated. Q.E.D. ■

Claim 3: *Suppose* $z_1 < z_1^E$*. If* R_1 *is strictly increasing in* i_2*, then there exists a (nondegenerate) set of album 1 qualities where Label* I *finances album 2.*

Proof. If Label I exercises its option on album 2, then its net profits in period 2 are

$$\pi^I(i_1, z_1) = E_{Z_2|Z_1}\left\{(1-\alpha)(R_1(i_1, i_2^I, z) + R_2(i_1, i_2^I, z))\right\} - i_2^I.$$

If Label I decides not to exercise its option on album 2, then its net profits in period 2 are

$$\pi^I(i_1, 0, z_1) = E_{Z_2|Z_1}(1-\alpha)(R_1(i_1, 0, z)).$$

Taking differences,

$$\begin{aligned} \pi^I(i_1, z_1) - \pi^I(i_1, 0, z_1) - F &= E_{Z_2|Z_1}(1-\alpha)[R_1(i_1, i_2^I, z) - R_1(i_1, 0, z)] \\ &\quad + [E_{Z_2|Z_1}(1-\alpha)R_2(i_1, i_2^I, z) - i_2^I - F]. \end{aligned}$$

Applying Claim 1, the first term in brackets is always positive and the second term is positive for $z_1 > z_1^E$. The result then follows from the fact that both terms are continuous in z_1 and the second term is strictly increasing in z_1. Q.E.D. ■

Claim 4: *Suppose $z_1 \geq z_1^E$. If R_1 is strictly increasing in i_2, then Label I always finances album 2.*

Proof. Assuming that Label E bids its willingness to pay for the rights to album 2, the artist's profits from switching to Label E is

$$d^A(i_1, i_2^E(i_1, z_1), z_1) = E_{Z_2|Z_1}\{[R_2(i_1, i_2^E(i_1, z_1); z) - i_2^E(i_1, z_1) - F] + \alpha R_1(i_1, i_2^E(i_1, z_1), z)\}.$$

The term in brackets is the amount that the artist can expect to earn in advances and royalties from label E; the second term is the royalties on album 1 that it obtains from Label I. If the artist switches labels, then Label I's payoff is

$$d^I(i_1, i_2^E(i_1, z_1), z_1) = E_{Z_2|Z_1}[(1-\alpha)(R_1(i_1, i_2^E(i_1, z_1), z)].$$

If Label L matches Label E's bid, then its payoff is

$$\begin{aligned} V(i_1, i_2^I(i_1, z_1), z_1) &= E_{Z_2|Z_1}\{(R_1(i_1, i_2^I(i_1, z_1); z) + R_2(i_1, i_2^I(i_1, z_1); z)\} \\ &\quad - i_2^I(i_1, z_1) - F - d^A(i_1, i_2^E(i_1, z_1), z_1). \end{aligned}$$

Here the bargaining surplus

$$S(i_1, z_1) = V(i_1, i_2^I(i_1, z_1), z_1) - d^I(i_1, i_1^E(i_1, z_1), z_1)$$

is strictly positive if R_1 is strictly increasing in z_1. Q.E.D. ■

Table 1: Summary Statistics

				Percentiles		
	N	Mean	Std. Dev.	.10	.50	.90
Date of release:						
album 1	355	13may1996	102	22aug1993	05may1996	28feb1999
2	355	20jul1998	108	23jul1995	02aug1998	27may2001
3	178	03jun1999	90	13oct1996	04aug1999	05aug2001
4	74	08jan2000	73	19apr1998	09feb2000	28oct2001
overall						
First year sales:						
album 1	355	312,074	755,251	7,381	78,360	781,801
2	355	367,103	935,912	10,705	55,675	951,956
3	178	450,716	867,630	7,837	71,674	1,461,214
4	74	316,335	579,869	6,137	87,898	912,078
overall	962	358,362	836,366	8,938	68,059	976,853
First 4 weeks / First year:						
album 1	355	.121	.111	.0161	.0846	.265
2	355	.263	.137	.0855	.263	.441
3	178	.305	.131	.134	.305	.5
4	74	.312	.144	.119	.294	.523
overall	962	.222	.15	.0341	.208	.431
Peak sales week:						
album 1	355	31.9	47.8	0	15	87
2	355	7.83	23.1	0	0	28
3	178	4.05	13.1	0	0	12
4	74	5.42	16.6	0	0	19
overall	962	15.8	35.3	0	1	44
Weeks between releases:						
1 & 2	355	114	53.5	58	107	179
2 & 3	178	111	46.7	58	104	169
3 & 4	74	93.1	36.8	50	88	154

Table 2: Seasonality in release dates

	Percent of releases occurring				
Month	Album 1 (n=355)	Album 2 (n=355)	Album 3 (n=178)	Album 4 (n=74)	Overall (n=962)
Jan	3.94	3.10	3.37	2.70	3.43
Feb	8.17	4.23	3.93	1.35	5.41
Mar	13.24	9.58	11.80	10.81	11.43
Apr	9.01	8.45	8.99	6.76	8.63
May	11.83	9.01	7.30	8.11	9.67
Jun	7.61	12.68	6.74	14.86	9.88
Jul	8.45	9.01	10.11	10.81	9.15
Aug	11.55	9.58	10.67	12.16	10.71
Sep	7.32	11.27	11.80	14.86	10.19
Oct	12.39	10.70	16.29	6.76	12.06
Nov	5.92	11.83	6.74	5.41	8.21
Dec	0.56	0.56	2.25	5.41	1.25

Table 3: Determinants of elapsed time between releases

	Elapsed time between:		
	1 and 2	2 and 3	3 and 4
First six months' sales	-0.006	-0.003	0.018
	(0.014)	(0.011)	(0.029)
Decline rate (prev. album)	-0.017	0.078	0.117
	(0.049)	(0.078)	(0.150)
Rap	-0.075	0.113	0.715
	(0.138)	(0.212)	(0.319)
Country	0.774	0.406	0.404
	(0.164)	(0.210)	(0.312)
Years since 1993	0.082	0.165	0.214
	(0.031)	(0.057)	(0.099)
N	355	177	74
log likelihood	-1715.07	-737.83	-243.06

Estimated coefficients from Cox proportional hazard models, with standard errors in parentheses. A positive coefficient means that an increase in the corresponding covariate is associated with an increased hazard rate (i.e., shorter time between releases). The estimation does *not* include right-censored observations—i.e., artists for whom the next album was not released before the end of our sample period.

Table 4: Estimated Effects of New Releases on Sales of Catalog Albums

Month (relative to release date)	Album pair					
	2→1	3→1	4→1	3→2	4→2	4→3
t=-13	-0.006	0.016	0.012	0.041	-0.005	-0.008
(0.017)	(0.027)	(0.048)	(0.025)	(0.041)	(0.041)	
t=-12	0.022	0.030	0.012	0.013	0.036	0.051
	(0.022)	(0.033)	(0.056)	(0.032)	(0.048)	(0.049)
t=-11	0.044	0.020	0.031	-0.048	-0.011	-0.012
	(0.025)	(0.036)	(0.058)	(0.035)	(0.050)	(0.053)
t=-10	0.059	0.031	0.100	0.024	0.047	0.044
	(0.028)	(0.038)	(0.058)	(0.037)	(0.051)	(0.055)
t=-9	0.066	0.023	0.061	0.052	0.062	0.068
	(0.029)	(0.038)	(0.059)	(0.039)	(0.051)	(0.056)
t=-8	0.078	0.037	0.041	0.055	0.038	0.086
	(0.030)	(0.039)	(0.060)	(0.040)	(0.052)	(0.056)
t=-7	0.124	0.035	0.010	0.074	0.008	0.050
	(0.031)	(0.039)	(0.060)	(0.040)	(0.051)	(0.056)
t=-6	0.148	0.077	0.113	0.090	0.073	0.071
	(0.031)	(0.040)	(0.059)	(0.041)	(0.052)	(0.057)
t=-5	0.201	0.127	0.155	0.121	0.082	0.079
	(0.032)	(0.040)	(0.061)	(0.041)	(0.052)	(0.057)
t=-4	0.260	0.142	0.167	0.177	0.148	0.121
	(0.032)	(0.040)	(0.059)	(0.042)	(0.052)	(0.058)
t=-3	0.301	0.178	0.141	0.242	0.153	0.146
	(0.033)	(0.040)	(0.059)	(0.042)	(0.053)	(0.058)
t=-2	0.346	0.240	0.197	0.257	0.207	0.242
	(0.033)	(0.040)	(0.060)	(0.042)	(0.053)	(0.058)
t=-1	0.419	0.290	0.299	0.332	0.280	0.231
	(0.033)	(0.041)	(0.062)	(0.042)	(0.053)	(0.059)
t=0	0.471	0.355	0.377	0.361	0.328	0.273
	(0.033)	(0.041)	(0.062)	(0.043)	(0.054)	(0.059)
t=1	0.449	0.267	0.323	0.311	0.275	0.290
	(0.034)	(0.041)	(0.062)	(0.043)	(0.054)	(0.059)
t=2	0.443	0.317	0.288	0.310	0.220	0.188
	(0.034)	(0.041)	(0.063)	(0.043)	(0.054)	(0.060)
t=3	0.425	0.311	0.279	0.286	0.231	0.247
	(0.034)	(0.041)	(0.062)	(0.043)	(0.054)	(0.060)
t=4	0.455	0.299	0.315	0.271	0.314	0.158
	(0.034)	(0.041)	(0.063)	(0.043)	(0.054)	(0.060)
t=5	0.455	0.285	0.297	0.254	0.226	0.252
	(0.034)	(0.042)	(0.062)	(0.044)	(0.054)	(0.061)
t=6	0.492	0.289	0.333	0.277	0.277	0.225
	(0.034)	(0.042)	(0.062)	(0.044)	(0.055)	(0.060)
t=7	0.509	0.296	0.298	0.263	0.259	0.189
	(0.035)	(0.042)	(0.063)	(0.044)	(0.055)	(0.061)
t=8	0.516	0.325	0.273	0.273	0.236	0.197
	(0.035)	(0.042)	(0.063)	(0.044)	(0.055)	(0.061)

(continued next page)

Table 4: (continued)

Month (relative to release date)	Album pair					
	2→1	3→1	4→1	3→2	4→2	4→3
t=9	0.474	0.324	0.260	0.268	0.237	0.195
	(0.035)	(0.042)	(0.063)	(0.044)	(0.056)	(0.061)
t=10	0.490	0.356	0.222	0.312	0.234	0.180
	(0.035)	(0.042)	(0.064)	(0.044)	(0.056)	(0.062)
t=11	0.489	0.337	0.178	0.339	0.154	0.155
	(0.035)	(0.042)	(0.063)	(0.045)	(0.056)	(0.062)
t=12	0.495	0.375	0.263	0.336	0.170	0.132
	(0.035)	(0.042)	(0.063)	(0.045)	(0.056)	(0.063)
t=13	0.530	0.323	0.293	0.289	0.180	0.173
	(0.035)	(0.042)	(0.065)	(0.045)	(0.056)	(0.063)
t=14	0.562	0.327	0.300	0.299	0.273	0.142
	(0.035)	(0.043)	(0.064)	(0.045)	(0.057)	(0.064)
t=15	0.530	0.245	0.301	0.255	0.261	0.219
	(0.036)	(0.043)	(0.064)	(0.045)	(0.057)	(0.064)
t=16	0.517	0.227	0.234	0.244	0.177	0.188
	(0.036)	(0.043)	(0.066)	(0.046)	(0.057)	(0.064)
t=17	0.533	0.233	0.298	0.213	0.217	0.144
	(0.036)	(0.043)	(0.065)	(0.046)	(0.057)	(0.064)
t=18	0.532	0.251	0.253	0.223	0.261	0.064
	(0.036)	(0.043)	(0.066)	(0.046)	(0.058)	(0.065)
t=19	0.545	0.196	0.404	0.161	0.275	0.231
	(0.036)	(0.043)	(0.067)	(0.046)	(0.058)	(0.065)
t=20	0.561	0.260	0.347	0.172	0.281	0.220
	(0.037)	(0.044)	(0.066)	(0.047)	(0.058)	(0.066)
t=21	0.515	0.250	0.269	0.178	0.209	0.222
	(0.037)	(0.043)	(0.066)	(0.047)	(0.057)	(0.066)
t=22	0.547	0.255	0.284	0.168	0.221	0.254
	(0.037)	(0.044)	(0.066)	(0.047)	(0.058)	(0.066)
t=23	0.561	0.226	0.267	0.183	0.156	0.139
	(0.037)	(0.044)	(0.066)	(0.047)	(0.058)	(0.067)
t=24	0.566	0.233	0.275	0.154	0.258	0.222
	(0.037)	(0.044)	(0.067)	(0.047)	(0.059)	(0.067)
t=25	0.581	0.251	0.271	0.137	0.223	0.179
	(0.037)	(0.044)	(0.067)	(0.047)	(0.059)	(0.068)
# albums	338	162	66	173	70	74
# observations	33,581	15,952	5,539	17,073	5,930	6,281
$\hat{\rho}$	.800	.659	.532	.736	.536	.637

Estimates of the regression described in equation 1, with standard errors in parentheses corrected for heteroskedasticity across albums and autocorrelation within albums. Estimated coefficients for time and seasonal dummies are suppressed to save space. Each column represents an album pair: e.g., the column labeled 4→2 lists the estimated effects of album 4's release on the sales of album 2. $t = 0$ is the first week following the release of the new album. The $\hat{\rho}$'s are the estimated AR(1) coefficients, reflecting the degree of serial correlation in demand shocks for a given album.

Table 5: Estimated effects of new releases: first-differenced model

Month (relative to release date)	Album pair					
	2→1	3→1	4→1	3→2	4→2	4→3
t=-13	-0.024	0.020	0.030	0.041	0.018	0.011
	(0.016)	(0.027)	(0.047)	(0.024)	(0.043)	(0.041)
t=-12	0.010	0.020	0.008	-0.029	0.048	0.051
	(0.016)	(0.027)	(0.049)	(0.024)	(0.044)	(0.041)
t=-11	0.012	-0.004	0.031	-0.060	-0.039	-0.056
	(0.017)	(0.027)	(0.051)	(0.024)	(0.044)	(0.041)
t=-10	0.010	0.018	0.078	0.073	0.060	0.058
	(0.016)	(0.027)	(0.048)	(0.024)	(0.044)	(0.041)
t=-9	-0.000	-0.002	-0.035	0.031	0.009	0.022
	(0.016)	(0.027)	(0.048)	(0.024)	(0.044)	(0.041)
t=-8	0.008	0.021	-0.011	0.011	-0.019	0.026
	(0.016)	(0.027)	(0.049)	(0.024)	(0.044)	(0.041)
t=-7	0.044	0.007	-0.017	0.029	-0.018	-0.017
	(0.017)	(0.027)	(0.050)	(0.024)	(0.043)	(0.040)
t=-6	0.022	0.054	0.112	0.028	0.080	0.037
	(0.017)	(0.028)	(0.048)	(0.024)	(0.044)	(0.039)
t=-5	0.054	0.051	0.038	0.038	0.017	0.018
	(0.016)	(0.027)	(0.047)	(0.024)	(0.044)	(0.041)
t=-4	0.057	0.025	0.020	0.060	0.061	0.042
	(0.016)	(0.027)	(0.047)	(0.024)	(0.044)	(0.041)
t=-3	0.042	0.043	-0.026	0.073	-0.001	0.023
	(0.016)	(0.027)	(0.047)	(0.024)	(0.044)	(0.041)
t=-2	0.050	0.077	0.060	0.022	0.055	0.092
	(0.016)	(0.027)	(0.047)	(0.025)	(0.044)	(0.041)
t=-1	0.079	0.064	0.101	0.089	0.080	-0.012
	(0.017)	(0.027)	(0.048)	(0.025)	(0.044)	(0.041)
t=0	0.055	0.075	0.088	0.040	0.060	0.045
	(0.017)	(0.027)	(0.050)	(0.024)	(0.044)	(0.041)
t=1	-0.018	-0.082	-0.053	-0.038	-0.047	0.015
	(0.016)	(0.027)	(0.048)	(0.025)	(0.044)	(0.041)
t=2	-0.007	0.047	-0.030	0.003	-0.057	-0.107
	(0.016)	(0.027)	(0.048)	(0.025)	(0.044)	(0.041)
t=3	-0.026	-0.010	-0.016	-0.023	0.011	0.057
	(0.016)	(0.027)	(0.048)	(0.025)	(0.044)	(0.041)
t=4	0.018	-0.013	0.030	-0.014	0.089	-0.085
	(0.016)	(0.027)	(0.048)	(0.025)	(0.044)	(0.040)
t=5	-0.019	-0.014	-0.003	-0.022	-0.077	0.102
	(0.016)	(0.027)	(0.047)	(0.024)	(0.044)	(0.040)
t=6	0.013	-0.010	0.023	0.019	0.048	-0.034
	(0.016)	(0.028)	(0.047)	(0.025)	(0.044)	(0.040)
t=7	-0.003	-0.007	-0.038	-0.021	-0.031	-0.050
	(0.017)	(0.028)	(0.047)	(0.025)	(0.044)	(0.040)
t=8	-0.008	0.020	-0.024	0.006	-0.018	0.004
	(0.016)	(0.028)	(0.050)	(0.025)	(0.044)	(0.041)

(continued next page)

Table 5: (continued)

Month (relative to release date)	Album pair					
	2→1	3→1	4→1	3→2	4→2	4→3
t=9	-0.050	-0.016	-0.010	-0.014	-0.003	-0.006
	(0.016)	(0.027)	(0.050)	(0.025)	(0.044)	(0.041)
t=10	0.014	0.011	-0.018	0.029	0.002	-0.007
	(0.017)	(0.027)	(0.050)	(0.025)	(0.044)	(0.041)
t=11	-0.003	-0.038	-0.015	0.015	-0.065	-0.007
	(0.016)	(0.028)	(0.047)	(0.025)	(0.044)	(0.041)
t=12	-0.007	0.018	0.091	-0.022	0.019	-0.017
	(0.017)	(0.028)	(0.047)	(0.025)	(0.044)	(0.041)
t=13	0.023	-0.059	0.017	-0.051	-0.000	0.029
	(0.017)	(0.028)	(0.047)	(0.025)	(0.044)	(0.041)
t=14	0.021	0.005	-0.001	0.015	0.074	-0.053
	(0.017)	(0.028)	(0.047)	(0.025)	(0.045)	(0.041)
t=15	-0.037	-0.063	-0.010	-0.027	-0.029	0.056
	(0.017)	(0.028)	(0.048)	(0.025)	(0.045)	(0.041)
t=16	-0.013	-0.002	-0.057	-0.002	-0.078	-0.035
	(0.016)	(0.028)	(0.048)	(0.025)	(0.045)	(0.041)
t=17	0.019	0.024	0.067	-0.017	0.049	-0.047
	(0.016)	(0.027)	(0.048)	(0.025)	(0.045)	(0.041)
t=18	-0.003	0.027	-0.035	0.013	0.064	-0.065
	(0.016)	(0.027)	(0.048)	(0.024)	(0.045)	(0.041)
t=19	0.007	-0.055	0.159	-0.060	0.027	0.176
	(0.016)	(0.027)	(0.050)	(0.025)	(0.045)	(0.042)
t=20	0.008	0.066	-0.049	0.014	0.020	-0.003
	(0.016)	(0.027)	(0.048)	(0.025)	(0.045)	(0.041)
t=21	-0.050	-0.014	-0.070	0.005	-0.070	0.004
	(0.017)	(0.027)	(0.048)	(0.025)	(0.042)	(0.041)
t=22	0.030	0.000	0.022	-0.009	0.021	0.029
	(0.016)	(0.027)	(0.048)	(0.024)	(0.042)	(0.040)
t=23	0.010	-0.026	-0.013	0.019	-0.066	-0.114
	(0.016)	(0.027)	(0.048)	(0.024)	(0.044)	(0.041)
t=24	-0.007	0.010	-0.004	-0.027	0.092	0.055
	(0.016)	(0.027)	(0.048)	(0.025)	(0.044)	(0.042)
t=25	0.001	0.020	-0.005	-0.013	-0.043	-0.067
	(0.017)	(0.028)	(0.048)	(0.025)	(0.044)	(0.042)
# albums	338	162	66	173	70	74
# observations	33,509	15,904	5,507	17,038	5,923	6,270

Estimates of the regression described in equation 2, with $\Delta ln(sales)$ as the dependent variable. Standard errors (in parentheses) corrected for heteroskedasticity across albums and autocorrelation within albums. Estimated coefficients for time and seasonal dummies are suppressed to save space.

Table 6: Externalities and hits

Album 1, Album 2:	Hit, Hit	Hit, Not	Not, Hit	Not, Not
N	53	45	34	206
Median # weeks to release 2	108	124	101	104
Median weekly sales (album 1) prior to release:	1,888	318	341	154
Median weekly decline around release:	-0.018	-0.016	0.009	-0.007
Estimated total change in sales:	23,766	679	30,292	954
Average of (first month sales)/(first year sales):	0.73	0.85	0.55	0.62
Album 2, Album 3:	Hit, Hit	Hit, Not	Not, Hit	Not, Not
N	49	13	12	99
Median # weeks to release 3	105	117	95	103
Median weekly sales (album 1) prior to release:	1,555	466	844	85
Median weekly decline around release:	-0.012	-0.026	0.004	-0.010
Estimated total change in sales:	19,884	1,110	20,788	687
Average of (first month sales)/(first year sales):	0.73	0.84	0.59	0.65

Hits are defined as albums that sold over 250,000 units nationally in the first year. Albums that didn't clear this threshold are the "Not" albums (i.e., not hits). The estimated total change in sales reflects the increase over the 39-week treatment window.

Table 7: (First month's sales)/(First year's sales): Summary statistics by hit pattern

				Percentiles		
	N	Mean	Std. Dev.	.10	.50	.90
Album 1:						
Hits	98	.100	.135	.005	.050	.345
Non-hits	240	.129	.100	.022	.104	.265
Album 2:						
Hits following hits:	53	.271	.158	.093	.219	.508
Non-hits following hits:	45	.329	.130	.154	.325	.464
Hits following non-hits:	34	.152	.150	.020	.090	.363
Non-hits following non-hits:	206	.265	.120	.119	.264	.422
Album 3:						
Hits following hits:	49	.329	.147	.137	.302	.541
Non-hits following hits:	13	.380	.094	.218	.403	.465
Hits following non-hits:	12	.184	.157	.028	.139	.333
Non-hits following non-hits:	99	.298	.114	.149	.298	.452

Table 8: Prices for catalog titles at the time of a new release

			Percentiles			
	N	mean	std. dev.	.10	.50	.90
New release titles	20	13.24	1.50	11.49	13.49	14.49
Catalog titles by artists with new releases	84	13.77	2.28	10.99	13.98	16.98
Catalog titles by artists without new releases	160	14.04	2.50	10.99	13.99	17.98

Based on prices at Amazon.com in February 2005.

Figure 1: Album sales paths for two examples

Figure 2: Distributions of Elapsed Time Between Releases

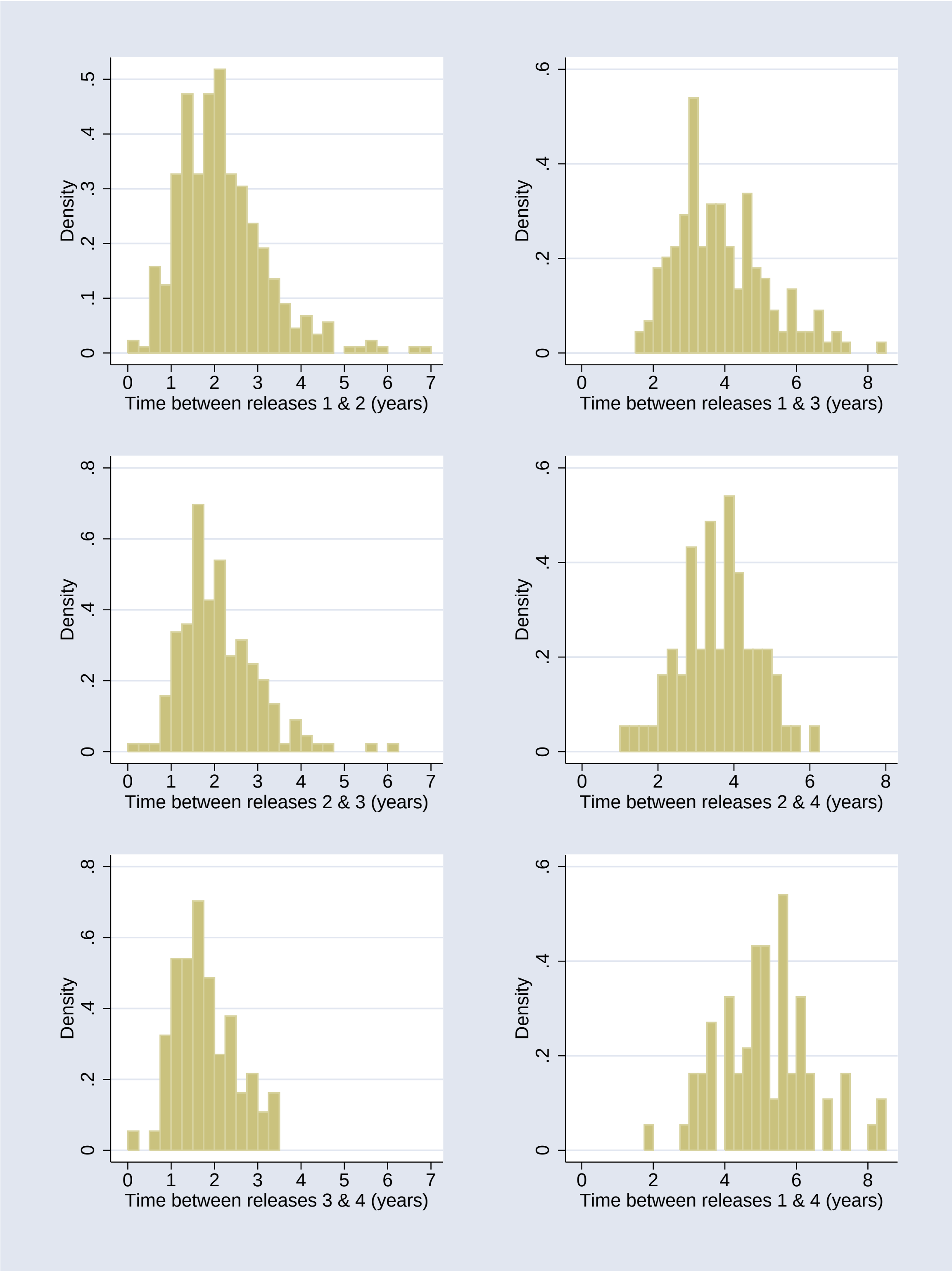

Figure 3: Time patterns of backward externalities

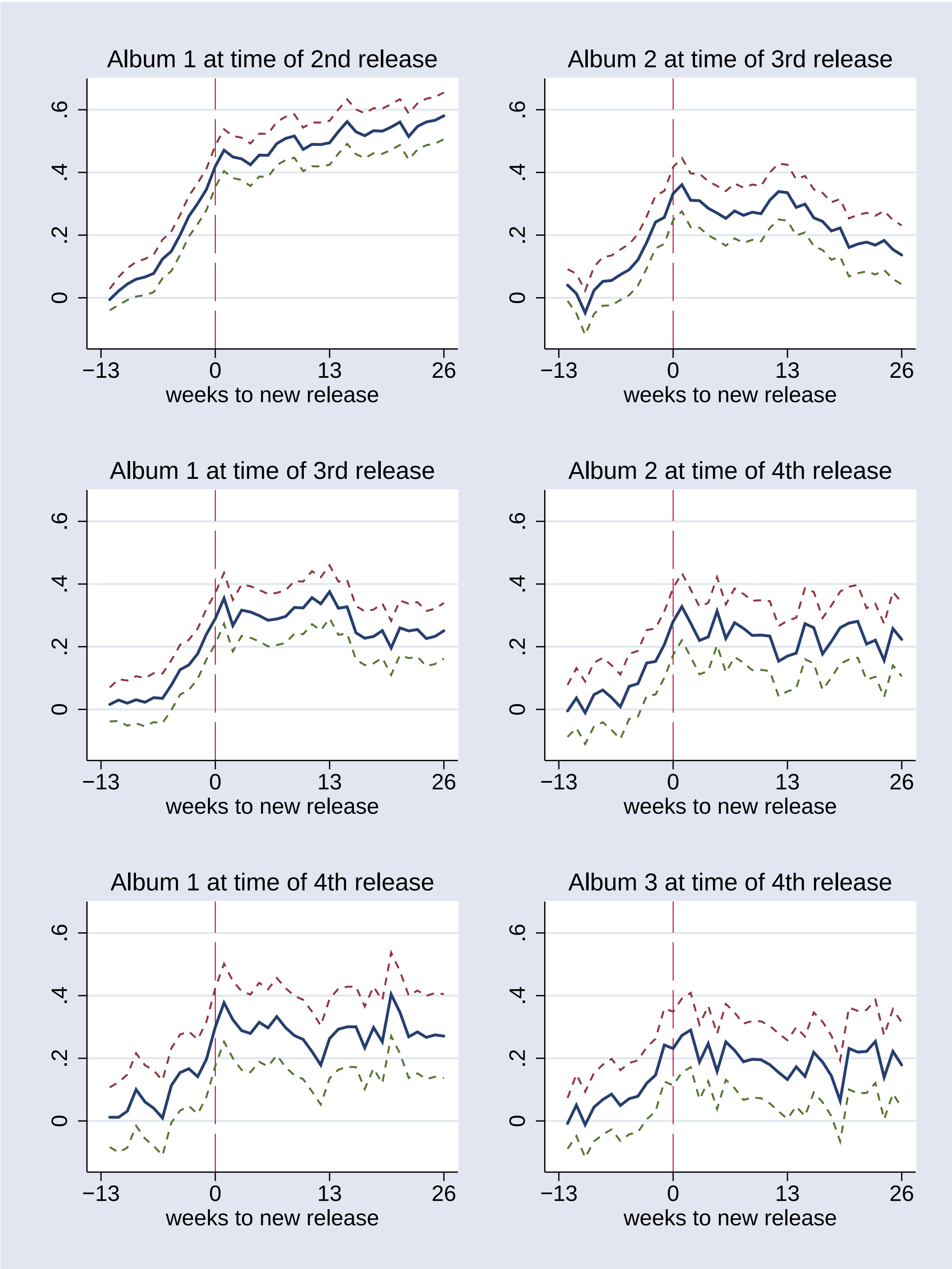

Figure 4: Time patterns of backward externalities: first-differences model