NBER WORKING PAPER SERIES

INFORMATION SPILLOVERS IN THE MARKET FOR RECORDED MUSIC

Ken Hendricks
Alan Sorensen

Working Paper 12263
http://www.nber.org/papers/w12263

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2006

We are grateful to seminar participants at Berkeley, Columbia, NYU, Penn, Penn State, Wisconsin, Yale, and the NBER Summer Institute for their comments and questions. We especially thank Dirk Bergmann, Greg Crawford, Marc Rysman, and Michael Whinston for suggestions that prompted changes in the paper, as well as Don Engel, Michael Lopez, and Ralph Peer for invaluable conversations that helped clarify several institutional details about the music industry. Chris Muratore and Nielsen SoundScan were very helpful in providing the data, and Natalie Chun and Abe Dunn provided outstanding research assistance. We are responsible for any errors or shortcomings that remain in spite of all the invaluable input. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

Information Spillovers in the Market for Recorded Music
Ken Hendricks and Alan Sorensen
NBER Working Paper No. 12263
May 2006
JEL No. D83, L15, L82

ABSTRACT

This paper studies the role of consumer learning in the demand for recorded music by examining the impact of an artist's new album on sales of past and future albums. Using detailed album sales data for a sample of 355 artists, we show that the release of a new album increases sales of old albums, and the increase is substantial and permanent—especially if the new release is a hit. Various patterns in the data suggest the source of the spillover is information: a new release causes some uninformed consumers to learn about their preferences for the artist's past albums. These information spillovers suggest that the high concentration of success across artists may partly result from a lack of information, and they have significant implications for investment and the structure of contracts between artists and record labels.

Ken Hendricks
Department of Economics
Bernard and Audre Rapoport Building
University of Texas at Austin
Austin, TX 78712
hendrick@eco.utexas.edu

Alan Sorensen
Graduate School of Business
Stanford University
518 Memorial Way
Stanford, CA 94305-5015
and NBER
asorensen@stanford.edu

1 Introduction

In entertainment industries such as movies, books, and music, many new products flow into the market each week. As a result, at any point in time, individual consumers may not know their preferences for many of the available products or even be aware of their existence. Consequently, market demand will depend not only on consumer preferences, but also on consumers' knowledge of the product space and the process by which they obtain this knowledge. We study this issue in the market for recorded music by measuring the impact of a new album release on sales of previous and future albums by the same artist. If consumer learning is important, then the promotional activity and radio airplay associated with a new release will enhance consumer awareness of the artist. Some of the newly informed consumers will want to buy the artist's old albums, leading to an increase in sales of these albums. We call this effect the *backward spillover*. The newly informed consumers also create a larger fan base for the artist's future albums, raising sales of these albums (relative to what they would have sold if the current album had never been released). We call this effect the *forward spillover*.

We find that the backward spillovers are on average positive, permanent, and both statistically and economically significant. Our empirical strategy for quantifying the average spillover is taken from the literature on treatment effects, but the effect is readily apparent from the album sales paths. Figure 1 shows two clear examples. 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 increased 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 was somewhat less dramatic, but still generated an increase in sales of the band's first album. In both examples, the backward spillover 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. Note that while the backward spillovers are clearly visible in the figure, forward spillovers are difficult to measure empirically. However, we show that under plausible conditions the backward spillover can be a good approximation to the forward spillover.

Our analysis of when and where the backward spillovers are large suggests that the main source of the spillovers is consumer learning. First, sales of catalog albums start to appear up to four weeks *prior* to the release of a new album, which we argue most likely reflects consumers learning about artists from pre-release radio airplay and other promotional activity. Second, the spillovers are larger when the new release is a hit, and especially large when the new release is a hit and the catalog album was not, which again is highly suggestive of consumers discovering artists who were previously unknown. Third, we show that backward spillovers are smaller in an artist's home market (i.e., the city where the artist began her career, and where there is presumably a larger stock of informed consumers), even though sales are on average higher in the home market. Taken together, these patterns suggest the backward spillover is mainly an information phenomenon: album releases generate new information, and this information leads some consumers to buy the artist's past albums.

This finding has broad implications for market outcomes. One distinguishing feature of entertainment industries like recorded music is that commercial success tends to be highly concentrated. Even among profitable albums, the distribution of returns is extremely skewed: a large share of total industry profit is claimed by a small number of very successful albums, and even fewer artists. The backward spillover suggests the correlation in consumer choices is not merely a reflection of the products' relative qualities. The skewness in the distribution of returns partly results from consumers' lack of information: consumers only learn about the most successful products, so success is self-reinforcing.

This implication is in line with herding models, such as those proposed by Banerjee [4] and Bikhchandani *et al* [8], but there is an important distinction in our case. The presence of uninformed consumers implies that many albums are undersold, but it does not necessarily imply that albums by hit artists are oversold. In the standard herding models, hit albums (and artists) would sell a lot fewer albums if consumers were fully informed and purchased albums on the basis of their idiosyncractic preferences rather than popularity. However, music differs from other herding markets such as books and restaurants. In the latter markets, consumers only observe what is popular; in music, successful albums get more radio airplay, generating signals that inform consumers about their own preferences for the album. Indeed, there is strong reason to believe that album success reinforces itself primarily through radio play: consumers buy only what they hear, and they only hear what others buy. If so, then the skewness from consumer learning arises not from hit albums being oversold, but rather from "good" albums being undersold because (for idiosyncratic reasons) they do not get very much airplay. This hypothesis also explains why catalog is promoted indirectly through new releases: many radio stations only play recently released albums.

Another distinguishing feature of entertainment industries is the aspect of joint production: in the music case, the spillovers occur within the context of a bilateral contracting problem, with the spillovers affecting the bargaining between artists and their record labels. The backward spillover generates a lock-in effect which keeps artists from switching labels, and the forward spillover generates a hold-up problem that can only be solved with long-term contracts. Thus, both types of spillover (backward and forward) have significant implications for investment and help explain the observed structure of contracts between artists and their record labels. Forward spillovers are also the central issue in legal disputes that arise when one label's artist is accused of using material from an album (not necessarily by that artist) owned by a different label.

We are not aware of prior empirical literature on information spillovers between products,[1] but our theoretical framework is similar to prior work on brand extension. Choi [10], Cabral [9], and Wernerfelt [22] have developed theoretical models that study the impact of information spillovers on firms' decisions about whether to release new products under existing brand names. 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 spillover), and the release of a high-quality new product can improve the brand image and boost sales of the existing product (the backward spillover).[2] There is a voluminous theoretical literature on the hold-up problem in contracts, but we are not aware of any that have studied the effect of backward spillovers.

The paper is organized as follows. In Section 2 we outline a simple model of consumer learning, give precise definitions of the backward and forward spillovers, and specify conditions under which the backward spillover is a good approximation for the forward spillover. Section 3 describes the data, which consist of weekly album sales histories for a sample of 355 artists. In Section 4 we describe the empirical strategy for estimating the backward spillover, which is taken from the literature on treatment effects. Essentially, the release of a new album is the treatment, and we

[1] Benkard's [7] study of learning by doing in aircraft production shows that learning spills over across aircraft types, but we have not seen any empirical papers that analyze information spillovers on the demand side of a market.

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

measure the treatment effect by comparing the sales paths of treated artists to those of a control group comprised of artists with the same number of catalog albums but who have not yet released new albums. We use fixed effects to control for time-invariant factors such as genre and artist popularity that may influence releases times, and we also estimate a first-differenced model that controls for possible correlation in the shape of the catalog album's sales path and release times. Section 5 reports the estimation results. Section 6 tests the model's predictions. In Section 7 we discuss the principal implications of our findings. Section 8 concludes.

2 Model

We introduce a two-period, two-album model of information spillovers to clarify the main ideas and frame the empirical analysis. In period 1, the artist signs a contract with a record label and produces an album of uncertain quality, which is denoted Z_1. (We denote random variables in upper case, and realizations in lower case.) The label observes an informative signal about Z_1 and, on the basis of this signal, decides whether to invest a fixed amount in marketing the album. If it does so, then the revenues generated by the album during period 1 are represented by R_{11}. In period 2, the label observes the realization r_{11} and decides whether to exercise its option to finance a second album. If it does so, the artist produces a second album of random quality Z_2. In period 2, revenues for albums 1 and 2 are R_{12} and R_{22}, respectively. The album qualities Z_1 and Z_2 are affiliated random variables: an artist that produces one high quality album is more likely to produce another high quality album. We assume that album price p is constant across time periods and across albums.

Consumers must learn about an album before making a decision to buy it. Survey evidence indicates that consumers learn their preferences about albums primarily by hearing them on the radio or by seeing music videos on television, and then sampling albums at listening posts in music stores.[3] Let X_1 denote the fraction of radio airplay that album 1 receives in period 1. We assume that X_1 is stochastically increasing in Z_1: a higher quality album is more likely to receive radio airplay. The probability a consumer learns about album 1 depends on the amount of airplay album

[3] In one national survey of music consumers conducted in 1994 [20] consumers were asked what motivated their recent music purchases, and the most common response was having heard the music on the radio. A more recent survey in 2006 [12] produced a similar finding: 55% of consumers said they learn about new music primarily from FM radio.

1 receives; here we simply assume that this probability equals x_1, the measure of its airplay. In period 2, X_2, the fraction of radio play given to album 2, is likely to depend in part on how well the first album did. Hence, we assume that X_2 is stochastically increasing in Z_1 and Z_2. Note that, given these assumptions, in both periods, a consumer is more likely to learn about a higher quality album. For simplicity, we will assume that album 1 receives no airplay in period 2.

Consumers' preferences are additive across albums, with the utility for album j net of price is given by

$$u_j = z_j - p + \varepsilon_j,$$

where ε_j is an idiosyncratic preference shock for album j. The preference shocks are i.i.d. across consumers, but may be correlated across albums for a given consumer. The joint distribution of $(\varepsilon_1, \varepsilon_2)$ is assumed to be symmetric with marginal distribution H. A consumer who has learned about album j buys it if u_j is positive.

We assume there is a continuum of consumers, and normalize the size of the market to one. Integrating over consumer demand, ex post revenues in period 1 are given by

$$r_{11} = px_1[1 - H(p - z_1)].$$

Album demand simply equals the fraction of consumers who are informed about the album and whose net utility from the album is positive.

In period 2, some consumers who hear the new album are discovering the artist for the first time. We assume that if a consumer learns about the artist in period 2, she learns about *both* albums—e.g., if she hears album 2 on the radio in period 2 and likes it, she will also check out the artist's previous albums when she visits the music store. Among these newly informed consumers, those with high enough valuations will purchase the old album, so album 1 revenues in period 2 are

$$r_{12} = p(1 - x_1)x_2[1 - H(p - z_1)].$$

Note that even if consumers "forget" about their preferences for album 1, the fraction of potential consumers for album 1 is given by the fraction that are learning about the album for the first time, because preferences do not change. We are also implicitly assuming that, in the absence of album 2, no additional consumers would discover album 1, and its sales would be zero in period 2.

What about sales of album 2? The probability that a consumer learns about album 2 is likely to depend on what he learned in period 1. The correlations in album qualities and in consumers' preferences give consumers an incentive to monitor the careers of artists they know, and especially those they like. This suggests consumers are more likely to learn about album 2 if they learned their preferences for album 1 in period 1. We make the simplifying assumption that if a consumer learned about album 1 in period 1, he will learn about album 2 with probability 1—i.e., consumers who discovered the artist at album 1 will be informed about album 2 when it is released. Consumers who did not discover the artist at album 1 will learn their preferences for album 2 with probability x_2. Ex post revenues for album 2 in period 2 are therefore given by

$$r_{22} = p(x_1 + (1 - x_1)x_2)[1 - H(p - z_2)].$$

Notice that our assumptions essentially imply that learning is about artists rather than albums. The fact that the release of album 2 causes some consumers to discover the artist and buy album 1 is what we call the *backward spillover*. The backward spillover is measured by r_{12}, since album 1 sales in the counterfactual world in which album 2 is not released are zero. In the empirical model described below in Section 4 we allow the counterfactual sales to be positive, but the conceptual framework is the same: we measure the backward spillover as the additional sales of album 1 that result directly from the release of album 2.

Notice also that consumers who buy the artist's debut album in period 1 constitute a "fan base" for the artist's second album. This is what generates the *forward spillover*, which we define as the difference between album 2 sales in the world where album 1 was released in period 1, versus sales in the counterfactual world where album 1 was never released. If X_2 is independent of Z_1 then this difference would be

$$p(1 - x_2)x_1[1 - H(p - z_2)].$$

The forward and backward spillovers therefore mirror each other: the forward spillover of a z quality album on a z' quality album is equal to the backward spillover of a z quality album on a z' quality album. This is important because, under the assumptions we have made here, even if forward spillovers cannot be measured empirically, their magnitudes can be inferred from the backward spillovers (which *can* be measured empirically). Moreover, if X_2 is stochastically increasing in Z_1, this would imply a forward spillover that is even larger than the backward spillover, since in that case album 1 generates both a fan base effect *and* an increased airplay effect for album 2.

To summarize, our simple model of spillovers is based on several key assumptions. First, prices are constant over time. We do not have any price data for the albums in our sample, so we cannot verify this assumption directly. However, we did collect price data for a sample of CDs offered at a major online retailer. Comparing prices for three groups of albums—new releases, catalog titles by artists with new releases, and catalog titles by artists without new releases—we found that although new releases tended to be discounted, the price distributions for the other two groups were indistinguishable. Catalog titles by artists who recently released a new album were no more likely to be discounted than other catalog titles. According to two retail store managers with whom we had conversations, even when catalog albums are discounted, the timing of the sales is not systematically related to new releases by the same artist. A second key assumption is that utility is additive (or more generally, submodular), which rules out complementarities in consumption.[4] We discuss the plausibility of a model with complementarities after reporting our results. Finally, the symmetry between forward and backward spillovers implicitly assumes that most of the variation in sales is due to variation in quality (i.e., Z) and not variation in promotional expenditures, which are not observable.

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

[4]Following the notation of the model, complementarities could be represented by having mean utility from purchasing *both* albums 1 and 2 be $z_1 + z_2 + \rho z_1 z_2 - 2p$, with $\rho > 0$ indexing the strength of the complementarity.

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.[5] 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 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 generate radio airplay and do not contain any new music that could be expected to affect demand for previous albums.[6] 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.[7] 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.[8] 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

[5]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.

[6]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.

[7]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.

[8]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.

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.

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 we can observe the flow of sales for prior albums at the time when a new album is released. Both cross-sectional and time-series variation can be exploited to measure the sales responses: for any given album, we can compare its sales path at the time of a new release to that album's sales history prior to the new release, and also to the sales paths of albums by other comparable artists who have not released new albums.

Table 1 summarizes various important aspects 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 are ten times sales at the median, and over 100 times sales 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 hit with her second album, 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.

Most albums' sales paths exhibit an early peak followed by a steady, roughly exponential decline. As indicated in the third and fourth panels of table 1, 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 increases 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 shows 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 for 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 clearly visible in the distribution of elapsed time between albums 3 and 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 come after January 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 spillover. Our approach to estimating the backward spillover is taken from the literature on treatment effects[9] 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.)

Without loss of generality, we focus on the first treatment episode: the release of album 2 and its impact on sales of album 1. 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. Time (t) is measured in terms of the number of periods since album 1 was released. Our objective is to estimate the average treatment effect on the treated (ATE) for each period of the treatment window. We focus on the treatment effect on the treated because of 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 is that, in each period, we observe only one outcome for each artist. The observed outcome for artist i in period t is

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

[9] See Wooldridge [23] for a summary.

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's 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.[10]

The regression model is as follows:

[10]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.

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

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).[11] 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 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 captured by 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 spillover reflects consumers' responses to these pre-release marketing campaigns. In some cases labels release singles from the new album in advance of the album itself, so that pre-release effects could also reflect advance airplay of the album's songs.[12] 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

[11]The results reported below are essentially unchanged if we control for seasonality with week-of-year dummies instead of month-of-year dummies.

[12]One might wonder whether the relevant event is the release of the single or the release of the album. Although we have data on when singles were released for *sale*, this does not correspond reliably with the timing of the release on the radio. Radio stations are given advance copies of albums to be played on the air, and a given single may be played on the radio long before it is released for sale in stores. Moreover, even when a single has been released in advance of the album, the label's promotional activity is still focused around the release date of the album.

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.

For the regression described above to yield consistent estimates of the treatment effect, the critical assumption 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) and of 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.[13] Whatever the reason for the relationship between the shape of the sales path and the time to the next release,

[13]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.

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 can 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 regressions in (1) and (2) separately for each of three "treatments:" the impact of the second, third, and fourth releases on sales of the previous album.[14] 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. 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

[14]Here we report results only for adjacent album pairs, but we have also measured the impact for non-adjacent pairs (e.g., the impact of album 3's release on sales of album 1). The effects for non-adjacent pairs are positive, statistically significant, and persistent, but slightly smaller than for adjacent album pairs.

four years after the second, then that artist is excluded from the regressions estimating the impact of album 3 on album 2.

Table 4 presents estimates of the regressions (1) and (2), 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 for specification (1) can be interpreted as approximate percentage changes in sales resulting from the new release, and for specification (2) they represent effects on the percentage rate of change in sales from week to week. The number of coefficients listed in Table 4 makes it somewhat difficult to read, so we summarize the results graphically in Figures 3 and 4. Figure 3 shows the estimated effects from specification (1), along with 95% confidence bands, for each 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 spillover is between albums 2 and 1, with estimates ranging between 40-55%. The spillovers for the remaining pairs of albums are smaller, ranging mostly between 15-30%. Figure 4 shows a comparison of the results from the two specifications. The solid line plots the cumulative impact implied by the estimated weekly coefficients from the first-differenced model (2), and the dashed line indicates the estimated effects from the levels regression (1). The implied effects are qualitatively and quantitatively very similar, which we interpret as reassuring evidence that our results are driven by real effects, not by subtle correlations between current sales flows and the timing of new releases.[15]

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 soon-to-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

[15]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. (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.) The estimates are always strongly significant.

of the release ($t = 0$ in the table), at which point there is a substantial spike upward in sales.

As mentioned previously, the pre-release effects most likely reflect promotional activity and radio airplay that occurs prior to the new album's release in stores. As a simple check, we estimated the regressions separately for artists who released singles in advance of the new album vs. artists who released singles after the release of the album. (Of the album releases represented in our sample, 23% were preceded by the release of a single, while another 16% had single releases occurring after the album release.) The pre-release effects on the catalog album's sales are much larger for the artists with advance singles,[16] which is consistent with the idea that these effects reflect consumer responses to the information flows associated with promotion and airplay.

The estimated effects are remarkably persistent: especially for the impact of album 2 on album 1, the spillovers do not appear to be transitory. 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.)

The main conclusion that we draw from the above results is that the backward spillover is on average significantly positive and permanent. There is no evidence of new albums cannibalizing sales of catalog albums or of new albums shifting demand for catalog albums from future periods towards the release period. The "buzz" and increased airplay around the time of a new release could accelerate the arrival of consumers at the store. 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 in sales would be transitory and eventually the spillover would become negative. We have tried longer treatment windows. In some cases, the treatment effect does die out eventually but in none of the cases does the treatment effect turn negative. Thus, the release of a new album generates permanent increases in demand for past albums, inducing purchases by customers who would not have otherwise purchased.

[16]For example, artists with pre-album singles had roughly 30-40% sales increases of album 1 in the three weeks prior to the release of album 2, while artists with post-album singles had increases of 5-20%.

6 Predictions

Our model predicts that the backward spillover will be smaller when more consumers know about the artist. We test this prediction in two ways. First, we examine how the backward spillover varies with the catalog album's success and the new album's success. Holding constant the success of the catalog album, the backward spillover should be larger if the new album is a hit, because uninformed consumers will be more likely to hear it, like it, and like the catalog album as well. It will be especially large when the catalog album was not a success because in this case the stock of uninformed consumers is quite large. Second, we examine geographic variation: i.e., instead of comparing national sales across artists, we can compare sales across markets for a given artist. An especially informative comparison is between an artist's home market (i.e., the city where the artist's career began) and other markets. Because new artists tend to have geographically limited concert tours—in many cases performing only in local clubs—artists in their early careers can be popular in their home markets while still relatively unknown on a national scale. If consumer learning is driving the backward spillovers, we should observe smaller spillovers in the artist's home market than in other markets.

6.1 Hits Versus Non-Hits

We split our sample depending on whether or not 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.[17] 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 spillovers for each of the four categories in Table 5.

The table is based on estimates of the regression model computed separately for each subgroup.[18] 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

[17] 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).

[18] 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 spillovers.) 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 2.1% per week. So we take a hypothetical album, with weekly sales beginning at 1,888 and declining at 2.1% per week, and apply the percentage increases implied by our estimated coefficients. The predicted total increase in sales over the 39-week period is 22,161, or roughly $350,000 in additional revenues (again using a rough price of $16 per unit).

The patterns in Table 5 are consistent with the predictions of our model. The backward spillover is always larger when the new album is a hit, whether the previous album was a non-hit or a hit. The largest percentage increase occurs when a non-hit album is followed by a hit: for an artist whose second album was her first hit, we estimate that weekly sales of her first album more than double when the new album is released. We interpret this as reflecting such artists' larger stock of uninformed consumers and higher quality second albums.[19] The smallest increase occurs when a hit is followed by a non-hit—i.e., when the stock of uninformed consumers is low and the quality of the new album is low.

The same patterns hold when we examine the impact of the third release on the sales of album 2. The spillovers 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 5 is that although on average (across all types of albums) the backward spillovers are of modest economic significance, they are in fact quite large for the artists that matter: those who have hits or have the potential to produce hits. This implies that if the artist's next release has the potential to be a hit, then the backward spillover 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 artist, then no spillover will be expected.

[19] Recall that due to the forward spillover, it takes a higher intrinsic quality to generate a hit following a non-hit than a hit following a hit.

Indirect evidence of forward spillovers can be seen in albums' decline rates. Table 7 reports the ratio of sales in the first month to sales in the first year of a new release for each of four different success patterns (two hits, a hit followed by a non-hit, etc.). The ratio is a simple measure of how front-loaded an album's sales are. Sales of new albums should be considerably more front-loaded when the previous album was a hit, because in that case 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. 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.

Recall that in section 2 we argued that if the variation in sales across albums is due mostly to quality rather than promotion—i.e., if an album's success depends on its Z, and not on variable investments made by the record label—then the backward spillover can be a lower bound on the forward spillover.[20] In terms of the binary classification, the forward spillover from a hit catalog album to a hit new album is approximately equal to the backward spillover from a hit new album to a hit catalog album; similarly, the forward spillover of a hit catalog album to a non-hit new album is approximately equal to the backward spillover from a hit new album to a non-hit catalog album. Table 5 indicates that the magnitudes of the backward spillovers in these two cases are substantial. We conclude that the forward spillovers in the associated cases are similar and hence also significant.

6.2 Home Versus Non-Home Markets

We were able to determine the city of origin for 325 of the 339 artists included in the regression analysis of Table 4; 268 of these artists originated in the U.S., so we can observe sales in the home

[20]We do not directly observe investment levels or album quality so we cannot separately measure the impact of these two inputs on sales. However, we believe that the long-term success of an album is plausibly more a function of its quality than of the label's marketing expenditures.

market and compare them to sales in other markets across the nation. SoundScan reports album sales separately for 100 Designated Market Areas (DMAs), each one corresponding to a major metropolitan area such as Los Angeles or Boston. We determined each artist's city of origin, and labeled the nearest DMA to be the artist's home market.[21] It is easy to verify that artists are indeed more popular in their home markets: over 80% of debut albums had disproportionately high sales in the artist's home market, meaning that the home market's share of national first-year sales was higher than the typical share for other artists of the same genre. On average, the home market's share of national sales was 8 percentage points larger than would have been predicted based on that market's share of overall sales within the artist's genre.

Are backward spillovers smaller in artists' home markets? Using the market-level data, we estimate a variant of the regression model in (1):

$$y_{imt} = \alpha_0 + \alpha_i + \sum_{g=1}^{4} \theta_{gm} G_i^g + \lambda_1 t + \lambda_2 t^2 + \psi H_{im} + \sum_{k=2}^{12} \delta_k D_{it}^k + \sum_{s=-13}^{26} I_{it}^s (\beta_s + \gamma H_{im}) + \epsilon_{imt} \quad (3)$$

where y_{imt} is log sales of artist i's album in market m in week t; G_i^g is a dummy equal to one if artist i is in genre g (so the θ_{gm}'s are market×genre fixed effects); the D_{it}^k's are month-of-year dummies, the I_{it}^s's are the treatment dummies, and H_{im} equals one if market m is artist i's home market. The key differences between this model and the one described in equation (1) are that (i) we use market-level sales data, and control for heterogeneity in sales across markets using market×genre fixed effects;[22] (ii) we measure whether sales are on average higher in the artist's home market (i.e., the parameter ψ); and (iii) we allow the spillover effects to differ for home markets vs. other markets (via the parameter γ).

Table 6 reports the key estimates for three album pairs. The estimates of ψ confirm that on average sales are much higher in an artist's home market than in other markets. For the debut album, the coefficient of 0.814 implies that sales are over twice as high in the home market than in other markets, other things being equal. Notably, the home market advantage is smaller for later albums, which

[21] Roughly 20% of the artists are solo artists, and for these we were only able to find the city of birth—which is not necessarily the city in which the artist first began performing. However, it is plausible that solo artists are more well-known in their birth cities than in other cities nationwide, even if they began their performing careers elsewhere. In any case, all of our analyses deliver the same conclusions if we exclude solo artists.

[22] Note that we can alternatively include market×artist fixed effects. Doing so means we cannot estimate ψ, the coefficient on H_{im}, because H_{im} is collinear with the market×artist effect for the home market. Adopting this specification yields results for all the other parameters that are virtually identical to those we report for the model with market×genre effects.

is consistent with the notion that awareness of the artist becomes less geographically concentrated as the the artist's career progresses.

In spite of the fact that artists' albums are on average more successful in their home markets, the backward spillovers are on average *smaller* in home markets. The estimates of γ are similar across the album pairs, indicating that backward spillovers are 10-14 percentage points smaller in an artist's home market than in other markets. We interpret this as evidence in support of learning models. Relative to other markets, home markets have smaller stocks of uninformed consumers—potential buyers are already familiar with the artist—so the new album generates fewer additional sales of the catalog title.

6.3 Alternative explanations

Another potential model of the backward spillovers is one in which consumers have supermodular preferences over albums. If owning one album by an artist increases the marginal utility from purchasing other albums by that artist, then the backward spillover can occur because some consumers who were previously not willing to buy catalog album are willing to do so when it is bundled with the new album. Like the learning model, complementarities in consumption would predict persistent backward spillovers, since in both cases the new release directly changes the probability of wanting to purchase the catalog album. The pre-release effects would have to be interpreted as purchases by consumers who anticipate buying the new album when it is released. That is, even though the benefits of joint consumption cannot be obtained until both albums are available, the consumer decides to buy the catalog album immediately to obtain the additional benefits of consuming the catalog album before the new release. Note that the complementarities could be interpreted as a characterization of fans: e.g., 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.[23]

The predictions of a preference complementarity model would depend on how the distribution of preferences in the population varies across markets. The fact that artists tend to be more successful in their home markets could be explained as a selection effect, but the magnitude seems implausibly

[23]Becker, Grossman, and Murphy [5] used a model with complementarities to describe cigarette addiction. Specifically, utility is $u(c_t, c_{t-1})$, with a positive cross-partial.

large. A similar issue arises when trying to explain why the backward spillover is smaller in an artist's home market. Even for big hits, sales of any single album reflect a small fraction of the number of potential consumers—i.e., those in the right tail of the distribution. This suggests that if sales of a given album are high, then backward spillovers for that album will also be high, because there will be a higher density of consumers with reservation prices near the margin. However, we observe the opposite when we examine artists' home markets: sales are higher in the home market, but spillovers are smaller. In order to explain this pattern, one would have to argue that preference complementarities operate differently in artists' home markets than in other markets, or that the distribution of preferences is such that high levels of sales can be consistent with a low density of consumers near the margin.[24]

We think such explanations are rather implausible, but we cannot rule them out entirely. A more direct way to check for supermodular preferences would be to see whether consumer purchases of catalog albums are typically bundled with purchases of new albums. Unfortunately, we do not have data on individual consumers' purchases, which could in principle allow us to quantify the relative importance of learning and complementarities. Other studies of consumer learning have used consumer-level data to draw sharper distinctions between competing hypotheses or to quantify the relative importance of alternative models. For example, Ackerberg [1] [2] used data on yogurt purchases by a panel of households in combination with data on advertising exposure to distinguish the informative vs. prestige effects of advertising. Crawford and Shum [11] use consumer panel data to estimate a structural model measuring the effects of uncertainty and learning on prescription choices and treatment outcomes for anti-ulcer drugs. They infer the importance of consumer learning essentially by comparing initial drug choices to drug choices made later on, in combination with data on patient recovery. An interesting similarity to our results is their finding that the drugs' market shares would be much less concentrated if patients were fully informed about their match values with the various drugs.

[24] As long as album purchases are tail events, almost any ordinary distribution (e.g., exponential, normal, lognormal) would imply the opposite: high levels of sales imply a high density of consumers near the margin.

7 Discussion

What does the backward spillover imply about the efficiency of consumer choices? In the context of our model, the backward spillover means that many albums are undersold, in the sense that many would-be buyers remain uninformed. This is especially true of debut albums in non-home markets. Of course, if an artist produces a hit album that gets a lot of radio airplay, then more consumers will learn about the artist and some of the lost catalog sales will be recovered. But the situation is quite different for artists who produce mediocre albums or albums that serve niche markets. Consumers have a more difficult time learning about these artists because their albums may get little or no airplay. Furthermore, their careers are more likely to be truncated due to low sales. For example, if an artist's first two albums are only moderately successful, her label may decline to produce any future albums—even though with full information the artist would eventually become a success.[25] Hence, market learning on non-hit artists will tend to be incomplete and their albums undersold.

On the other hand, hit albums are unlikely to be oversold. In the standard herding model, consumers rationally ignore their own information to follow the herd: they buy what others buy. If consumers behave this way in the market for music, it would imply that not only are some albums undersold, but others are oversold. However, the latter phenomenon may be less relevant for recorded music. In markets such as restaurants and books, potential consumers only observe what other consumers buy, so they draw inferences about a product's quality only from the knowledge of its overall popularity. But, in music, when other consumers buy an album, the songs on that album get played more frequently on the radio, generating signals that inform consumers' about their *own* preferences for the album. As a result, consumers are less likely to herd on a bad album.[26] If, in addition, preferences are additive across albums by the same artist and across artists, then there is no meaningful substitution between albums. Artists may sell fewer albums than they would in the but-for world of complete information, but they do not lose sales to other albums.

The broad implication is that the distribution of sales in music is significantly more concentrated than it would be in the but-for-world of complete information. In other words, the high concentration of returns across artists partly reflects the way in which consumers learn about their

[25] We do not mean to suggest that all unsuccessful artists are potential stars, but rather that some potential stars' careers may be truncated because consumers were unaware of their music.

[26] This may partly explain why book sales are much more skewed than music sales. (See Sorensen [19] for some evidence and discussion of the skewed distribution of sales for hardcover fiction.)

preferences rather than their actual preferences.[27] Indeed, we suspect that the skewness in album sales largely reflects the skewness in radio airplay. If consumers in music markets buy only what they hear, and they hear only what others buy, then success reinforces itself, which generates a more skewed distribution of sales.[28] An interesting implication is that lower costs of acquiring information about albums (e.g., due to sharing and sampling of music on the internet) should make it easier for artists to become known, and since this has a disproportionate impact on lesser-known (i.e., less successful) artists, it would tend to make sales less concentrated across artists.

What does the backward spillover imply about investment in artists? The standard recording artist contract is a "work-for-hire" agreement, so the label owns the recording rights on albums.[29] The label pays the artist an advance to cover studio costs and "living costs," and the artist agrees to produce an album. After the artist delivers the album, the label decides whether to release it and how much to invest in marketing and promotion.[30] The investment decision is contingent on private information that is not verifiable and hence not contractible. The backward spillover implies that in order to ensure that investments in the new album are efficient, the rights on catalog albums have to be bundled with the rights on the new album. Otherwise, the label that owns the recording rights to the new album will not internalize the impact of its investment on sales of catalog. The incumbent label, who owns the recording rights to catalog, will have an advantage in bidding for the rights to the new album: it is willing to invest more and pay more for those rights than an outside label. Thus, the backward spillover tends to lock in the artist. However, it can be shown that the backward spillover by itself does not lead to inefficiencies in investment, nor does it necessitate long-term contracts.[31] Intuitively, when the label invests in the artist's debut album (and every subsequent album), it anticipates winning the rights to future albums and earning the rents associated with the

[27] Herding and other information-based models are not the only models that predict endogenously skewed market returns. For example, Becker and Murphy [6] propose a class of models in which social multiplier effects arise because a product's popularity increases the consumer's marginal utility of consumption, due perhaps to network effects in consumption. Such models cannot easily explain the patterns we observe in the backward spillovers, but they may be an additional source of skewness in the demand for music.

[28] Sorensen [19] documents a similar phenomenon of success breeding success in the market for books: appearing on the *New York Times* bestseller list has a direct positive impact on sales. The effect is negligible for well-known authors, but quite large for debut authors—a result that parallels the finding here that backward spillovers are largest when a hit new album is released by a relatively unknown artist.

[29] See Krasilovsky et al's [16] book on contracts in the music industry for more details. In some rare cases the artists negotiate reversions—i.e., ownership of the recording rights reverts to the artist after some period of time.

[30] Labels typically spend between \$250,000 to \$500,000 on marketing and promotion. The typical advance for a debut album is usually around \$150,000.

[31] In a previous version of this paper we described a formal model of contracting. It is available from the authors upon request.

backward spillover.

By giving the incumbent label a bidding advantage, the backward spillover may present a barrier to entry. Entrant labels cannot internalize the spillover, so they are at a disadvantage whenever the spillover 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 spillover by purchasing the rights to the artist's catalog from the incumbent label. Major distributors such as Sony, Time-Warner, BMG, Universal, and EMI adopted this strategy when they decided to vertically integrate backwards into the production of music: instead of bidding for the rights to new albums, they effectively purchased the recording rights to catalog by buying labels. However, as a matter of practice, the major distributors never sell catalog rights. The reasons for this are not entirely clear, but we suspect that frictions such as information asymmetries and strategic concerns tend to prevent such trades. The incumbent label has private information about the artist, so the usual adverse selection problems will inhibit the trading of artist's catalogs; and 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.

The forward spillover implies that the recording rights to the new album and future albums need to be bundled when the new album is released. Investments in new albums yield returns on future albums, both from the fan base generated by the first album (i.e., consumer learning), and from the information about the artist's quality that is generated by the first album (i.e., firm learning). If the album rights are not bundled, these returns will not be fully captured by the investing label: other labels can free-ride and selectively bid for new albums by artists whose previous albums did well. Hence, in the absence of a long-term contract, the artist will be able to capture some of those investment returns. This is the familiar holdup problem. It reduces the willingness of the label to invest in a new album, leading to underinvestment (and possibly no investment) in that album. Long-term contracts resolve the holdup problem. Our estimates of the backward spillover imply that forward spillovers are important, and consequently that the holdup problem is substantial. This helps explain why virtually all contracts between artists and labels are initially long-term contracts.

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 would likely lead to significant inefficiencies. Fewer albums would be produced, and a higher proportion of the albums would be by established artists.

In practice, artists may not be able to commit to a long-term contract. Contract terms are almost always renegotiated after an artist has a successful album. Artists gain bargaining leverage following a hit album because outside labels are willing to pay for the next album. The artist can exploit this leverage by strategically withholding or delaying new recordings, or by (with the help of a lawyer) getting out of a recording contract.[32] The outcomes of the renegotiations provide further evidence of the importance of the backward spillover. Artists almost always stay with their incumbent labels. In our sample, fewer than 10% of artists ever switched between major labels, and most of the observed switches were due to termination by the incumbent label. Furthermore, artists who negotiate "reversions" in their initial contracts—i.e., clauses stipulating that the rights to the masters revert to the artist after some number of years—typically lease their catalogs to the record label that is producing and distributing their current albums. Note that even when the catalog albums did poorly, and the artist has no bargaining leverage, the incumbent label still often exercises its option to produce another album. Backward spillovers help rationalize this fact: the incumbent label may still find it worthwhile to invest in another album (even though the outside label does not) because of the spillover onto catalog.

The structure of the contract also suggests that the parties anticipate the lack of commitment. If the artist could commit, the parties should sign a long-term contract in which the label is the residual claimant to album revenues, and the artist receives an up-front fee for each album, with the fee being contingent on the success of previous albums. In fact, the standard contract gives the artist royalties (usually between 10-14%) but requires her to repay advances out of royalties. The royalties are a potential source of inefficiency, since a label will underinvest if it has to share album revenues with the artist. However, artists almost never repay their accumulated advances to the labels—so in practice the royalties are not a problem.[33] We conjecture that the repayment

[32]Our description of contracting practices in this section is based largely on conversations with Don Engel, one of the more successful lawyers who specializes in renegotiating contracts. His press pseudonym is "Busta Contract."

[33]Recoupment of advances means that, for example, if the royalty rate is 12% and the advance is $200,000, the album has to generate 1.67 million dollars in revenue (approximately 100,000 CD sales) before the label shares revenues with the artist. If an artist's advances from previous albums are still unrecouped from her royalty revenues on that album, then the remaining debt (plus any additional advances) is recoverable from royalties on the new album.

of advances is due to the artist's inability to commit not to hold up the label by renegotiating her contract. By promising ex ante to repay some fraction of the advance from revenues of the current and future albums, the parties are able to write a contract in which ex post transfers between the artist and label are contingent on album sales, which are verifiable and hence contractible.

8 Conclusion

We find that the backward spillover is on average positive, substantial, and permanent. The magnitude of the backward spillover is strongly related to the relative success of the catalog and new release—the largest spillovers occur when the new album is the artist's first hit. The incidence of the backward spillover also varies geographically: although artists sell roughly twice as many albums in their home markets as in other markets, the backward spillover is significantly smaller in home markets. These findings strongly suggest that new albums are informative events, creating consumer awareness and generating signals that cause many consumers to re-evaluate their prior decisions not to buy catalog albums.

The presence of uninformed consumers implies a role for advertising. Indeed, Goeree [15] shows that advertising is important in the market for PCs, where she argues that the rapid pace of technological change leads consumers to be less than fully informed about the set of available products. Similarly, Ackerberg's studies of yogurt purchases suggest that advertising's influence stems largely from its ability to inform "inexperienced" consumers. In the present context, our interpretation of the backward spillovers raises the possibility that similar sales increases might be obtained through direct advertising—i.e., with respect to promoting catalog sales, marketing expenditures may be a substitute for new album releases. However, it is unlikely that spillover sales could be generated from advertising alone. The reason is that radio airplay seems to be the most important form of promotion, and the primary channel through which consumers learn about new music. This means that record labels cannot simply buy consumer awareness—at least not legally. It also implies that skewness in music sales partly reflects skewness in airplay. Verifying this hypothesis with data on airplay is an exercise we leave for future research.

Since successful artists usually obtain higher advances, over 95% of recording artists are unrecouped, so the label effectively earns 100% of the marginal revenues from album sales.

References

[1] Ackerberg, D., "Empirically Distinguishing Informative and Prestige Effects of Advertising," *RAND Journal of Economics*, vol. 32 no. 2, Summer 2001.

[2] Ackerberg, D., "Advertising, Learning, and Consumer Choice in Experience Good Markets: An Empirical Examination," *International Economic Review*, vol. 44 no. 3, August 2003.

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

[4] Banerjee, A. "A Simple Model of Herd Behavior," *Quarterly Journal of Economics*, v. 107 n. 3, August 1992, pp. 797-817.

[5] Becker, G., Grossman, M., and K. Murphy, "An Empirical Analysis of Cigarette Addiction," *American Economic Review*, v. 84 n. 3, June 1994, pp. 396-418.

[6] Becker, G., and Murphy, K., *Social Economics: Market Behavior in a Social Environment*, Harvard University Press, 2000.

[7] Benkard, C. L., "Learning and Forgetting: The Dynamics of Aircraft Production," *American Economic Review*, v. 90 n. 4, September 2000, pp. 1034-1054.

[8] Bikhchandani, S., Hirshleifer, D. and Welch, I., "A Theory of Fads, Fashions, Custom, and Cultural Change as Informational Cascades," *Journal of Political Economy*, 100, October 1992, pp. 992-1026.

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

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

[11] Crawford, G. and M. Shum, "Uncertainty and Learning in Pharmaceutical Demand," *Econometrica*, vol. 73, no. 4, July 2005.

[12] "Details About the AP/Rolling Stone Poll on Downloading Music," Associated Press, Feb. 2., 2006.

[13] Gandal, N., Sarit Markovich, Michael Riordan, "Ain't it 'Suite': Strategic Bundling in the PC Office Software Market," working paper, April 2004.

[14] Gentzkow, M., "Valuing New Goods in a Model with Complementarity: Online Newspapers," working paper, 2004.

[15] Goeree, M. S., 2005, "Advertising in the U.S. Personal Computer Industry," Claremont McKenna working paper, 2005.

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

[17] McMillan, J., "Trench Town Rock: The Creation of Jamaica's Music Industry," mimeo, June, 2005.

[18] Smith, L. and P. Sørensen, "Pathological Outcomes of Observational Learning," *Econometrica*, v. 68 n. 2, March 2000, pp. 371-398.

[19] Sorensen, A., "Bestseller Lists and Product Variety: The Case of Book Sales," Stanford University working paper, 2005.

[20] *Soundata National Music Consumer Study*, National Association of Recording Merchandisers, 1995.

[21] Terviö, M., "Mediocrity in Talent Markets," University of Berkeley working paper, December, 2004.

[22] 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.

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

Table 1: Summary Statistics

	N	Mean	Std. Dev.	Percentiles .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

Week (relative to release date)	Baseline model (1)			First-differenced model (2)		
	2→1	3→2	4→3	2→1	3→2	4→3
t=-13	-0.006	0.041	-0.008	-0.024	0.041	0.011
	(0.017)	(0.025)	(0.041)	(0.016)	(0.024)	(0.041)
t=-12	0.022	0.013	0.051	0.010	-0.029	0.051
	(0.022)	(0.032)	(0.049)	(0.016)	(0.024)	(0.041)
t=-11	0.044	-0.048	-0.012	0.012	-0.060	-0.056
	(0.025)	(0.035)	(0.053)	(0.017)	(0.024)	(0.041)
t=-10	0.059	0.024	0.044	0.010	0.073	0.058
	(0.028)	(0.037)	(0.055)	(0.016)	(0.024)	(0.041)
t=-9	0.066	0.052	0.068	-0.000	0.031	0.022
	(0.029)	(0.039)	(0.056)	(0.016)	(0.024)	(0.041)
t=-8	0.078	0.055	0.086	0.008	0.011	0.026
	(0.030)	(0.040)	(0.056)	(0.016)	(0.024)	(0.041)
t=-7	0.124	0.074	0.050	0.044	0.029	-0.017
	(0.031)	(0.040)	(0.056)	(0.017)	(0.024)	(0.040)
t=-6	0.148	0.090	0.071	0.022	0.028	0.037
	(0.031)	(0.041)	(0.057)	(0.017)	(0.024)	(0.039)
t=-5	0.201	0.121	0.079	0.054	0.038	0.018
	(0.032)	(0.041)	(0.057)	(0.016)	(0.024)	(0.041)
t=-4	0.260	0.177	0.121	0.057	0.060	0.042
	(0.032)	(0.042)	(0.058)	(0.016)	(0.024)	(0.041)
t=-3	0.301	0.242	0.146	0.042	0.073	0.023
	(0.033)	(0.042)	(0.058)	(0.016)	(0.024)	(0.041)
t=-2	0.346	0.257	0.242	0.050	0.022	0.092
	(0.033)	(0.042)	(0.058)	(0.016)	(0.025)	(0.041)
t=-1	0.419	0.332	0.231	0.079	0.089	-0.012
	(0.033)	(0.042)	(0.059)	(0.017)	(0.025)	(0.041)
t=0	0.471	0.361	0.273	0.055	0.040	0.045
	(0.033)	(0.043)	(0.059)	(0.017)	(0.024)	(0.041)
t=1	0.449	0.311	0.290	-0.018	-0.038	0.015
	(0.034)	(0.043)	(0.059)	(0.016)	(0.025)	(0.041)
t=2	0.443	0.310	0.188	-0.007	0.003	-0.107
	(0.034)	(0.043)	(0.060)	(0.016)	(0.025)	(0.041)
t=3	0.425	0.286	0.247	-0.026	-0.023	0.057
	(0.034)	(0.043)	(0.060)	(0.016)	(0.025)	(0.041)
t=4	0.455	0.271	0.158	0.018	-0.014	-0.085
	(0.034)	(0.043)	(0.060)	(0.016)	(0.025)	(0.040)
t=5	0.455	0.254	0.252	-0.019	-0.022	0.102
	(0.034)	(0.044)	(0.061)	(0.016)	(0.024)	(0.040)
t=6	0.492	0.277	0.225	0.013	0.019	-0.034
	(0.034)	(0.044)	(0.060)	(0.016)	(0.025)	(0.040)
t=7	0.509	0.263	0.189	-0.003	-0.021	-0.050
	(0.035)	(0.044)	(0.061)	(0.017)	(0.025)	(0.040)

(continued next page)

Table 4: (continued)

Week (relative to release date)	Baseline model (1) 2→1	3→2	4→3	First-differenced model (2) 2→1	3→2	4→3
t=8	0.516	0.273	0.197	-0.008	0.006	0.004
	(0.035)	(0.044)	(0.061)	(0.016)	(0.025)	(0.041)
t=9	0.474	0.268	0.195	-0.050	-0.014	-0.006
	(0.035)	(0.044)	(0.061)	(0.016)	(0.025)	(0.041)
t=10	0.490	0.312	0.180	0.014	0.029	-0.007
	(0.035)	(0.044)	(0.062)	(0.017)	(0.025)	(0.041)
t=11	0.489	0.339	0.155	-0.003	0.015	-0.007
	(0.035)	(0.045)	(0.062)	(0.016)	(0.025)	(0.041)
t=12	0.495	0.336	0.132	-0.007	-0.022	-0.017
	(0.035)	(0.045)	(0.063)	(0.017)	(0.025)	(0.041)
t=13	0.530	0.289	0.173	0.023	-0.051	0.029
	(0.035)	(0.045)	(0.063)	(0.017)	(0.025)	(0.041)
t=14	0.562	0.299	0.142	0.021	0.015	-0.053
	(0.035)	(0.045)	(0.064)	(0.017)	(0.025)	(0.041)
t=15	0.530	0.255	0.219	-0.037	-0.027	0.056
	(0.036)	(0.045)	(0.064)	(0.017)	(0.025)	(0.041)
t=16	0.517	0.244	0.188	-0.013	-0.002	-0.035
	(0.036)	(0.046)	(0.064)	(0.016)	(0.025)	(0.041)
t=17	0.533	0.213	0.144	0.019	-0.017	-0.047
	(0.036)	(0.046)	(0.064)	(0.016)	(0.025)	(0.041)
t=18	0.532	0.223	0.064	-0.003	0.013	-0.065
	(0.036)	(0.046)	(0.065)	(0.016)	(0.024)	(0.041)
t=19	0.545	0.161	0.231	0.007	-0.060	0.176
	(0.036)	(0.046)	(0.065)	(0.016)	(0.025)	(0.042)
t=20	0.561	0.172	0.220	0.008	0.014	-0.003
	(0.037)	(0.047)	(0.066)	(0.016)	(0.025)	(0.041)
t=21	0.515	0.178	0.222	-0.050	0.005	0.004
	(0.037)	(0.047)	(0.066)	(0.017)	(0.025)	(0.041)
t=22	0.547	0.168	0.254	0.030	-0.009	0.029
	(0.037)	(0.047)	(0.066)	(0.016)	(0.024)	(0.040)
t=23	0.561	0.183	0.139	0.010	0.019	-0.114
	(0.037)	(0.047)	(0.067)	(0.016)	(0.024)	(0.041)
t=24	0.566	0.154	0.222	-0.007	-0.027	0.055
	(0.037)	(0.047)	(0.067)	(0.016)	(0.025)	(0.042)
t=25	0.581	0.137	0.179	0.001	-0.013	-0.067
	(0.037)	(0.047)	(0.068)	(0.017)	(0.025)	(0.042)
# albums	338	173	74	338	173	74
# observations	33,581	17,073	6,281	33,509	17,038	6,270
$\hat{\rho}$	.800	.736	.637	-.220	-.270	-.266

Estimates of the regressions described in equations 1 and 2, 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 3→2 lists the estimated effects of album 3'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: Spillovers 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	342	154
Median weekly decline around release:	-0.021	-0.018	-0.018	-0.011
Estimated total change in sales:	22,161	660	14,557	883
Percentage change in sales:	42.7	7.2	148.5	17.6
Average of (sales before next release)/(first 4 years' 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:	-0.013	-0.026	0.004	-0.010
Estimated total change in sales:	19,884	1,110	20,788	687
Percentage change in sales:	40.6	9.5	56.4	24.6
Average of (sales before next release)/(first 4 years' 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 changes and percentage changes in sales reflect increases over the 39-week treatment window.

Table 6: Sales and spillovers in the artist's home market

	2→1	3→2	4→3
Home market ($\hat{\psi}$)	0.814	0.647	0.689
	(0.006)	(0.008)	(0.013)
Home market × new release period ($\hat{\gamma}$)	-0.105	-0.104	-0.137
	(0.010)	(0.013)	(0.018)
# observations	2,727,890	1,437,340	536,400
# artists	268	142	63

Estimates of the regression model described in equation (3); the dependent variable is log sales. $\hat{\psi}$ measures the average difference in log sales between the artist's home market vs. other markets, and $\hat{\gamma}$ measures the average difference in the backward spillover in the artist's home market vs. other markets. Other coefficients are omitted to save space.

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

Figure 1: Album sales paths for two examples

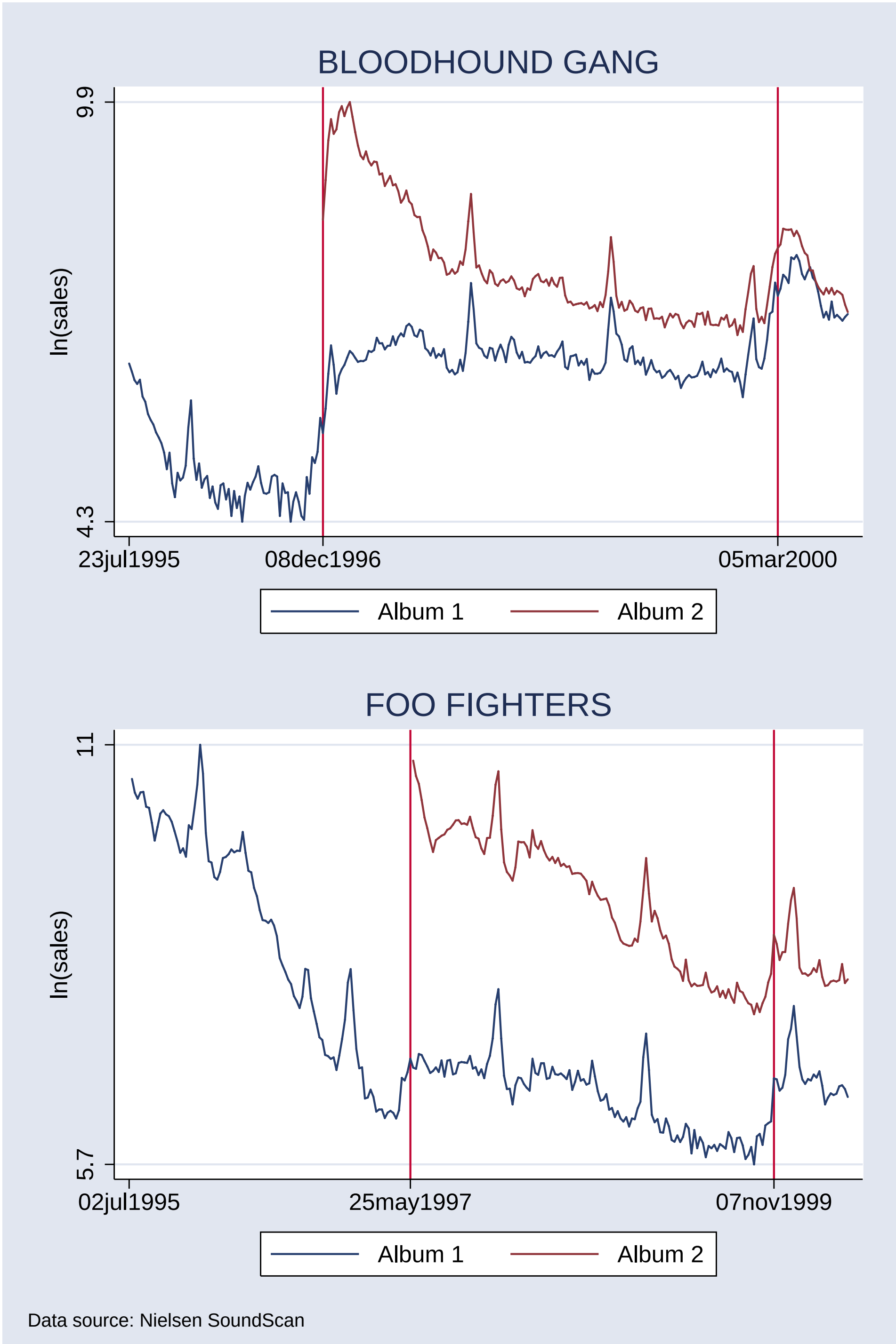

Figure 2: Distributions of Elapsed Time Between Releases

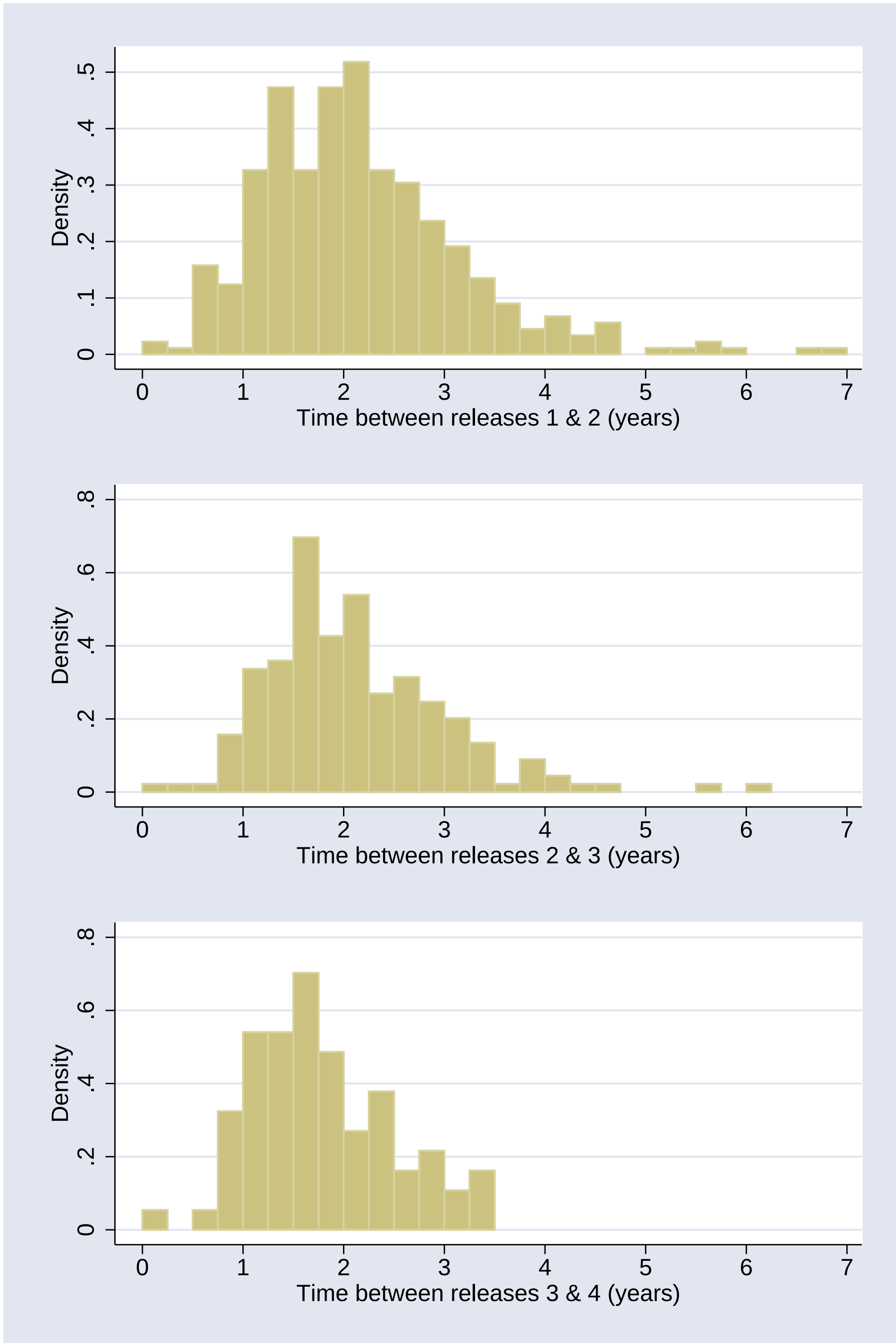

Figure 3: Time patterns of backward spillovers

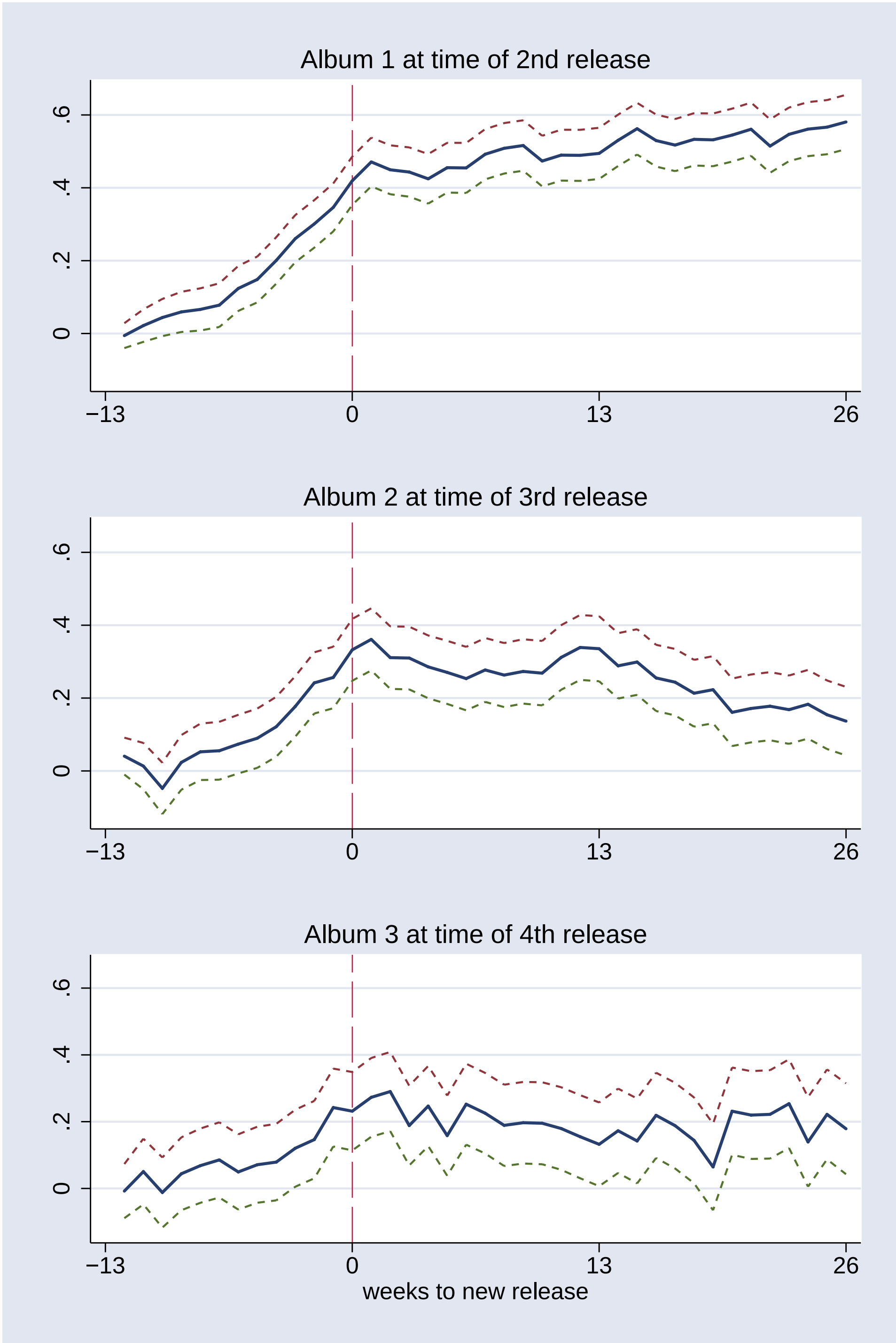

Figure 4: Time patterns of backward spillovers: first-differences model