An Empirical Study of the Impact of New Album Releases on Sales of Old Albums by the Same Recording Artist

Ken Hendricks
Department of Economics
Princeton University
University of Texas

Alan Sorensen
Graduate School of Business
Stanford University,

November 30, 2004
PRELIMINARY AND INCOMPLETE
Please do not quote

Abstract

This paper examines the impact of the release of a new album on sales of old albums by the same recording artist. We find that a new album increases sales of the old albums, and the increase is substantial and permanent. The pattern of spillovers across pairs of albums suggests that the main source of the spillover is consumers who change their minds about the value of buying old albums, either because of new information or because of preference complementarities with the new album. Our findings have important implications for contractual relationships between distributor and artist, for investment, marketing, and pricing of albums, and for market structure.

1 Introduction

In entertainment industries such as movies, books and music, new products flow into the market each month. The new products compete against each other as well as older products, often produced by the same suppliers. In this paper, we focus on the music industry and investigate the impact of a new album on sales of the artist's old albums. We distinguish two kinds of spillovers. The *forward spillover* refers to the impact of a new album on sales of future albums. At any point in time, consumers may be familiar with only a relatively small fraction of existing artists and to have heard only a small subset of the available albums. Since new releases typically get more playing time at radio stations than older releases, and artists frequently go on tour to promote a new album, the new release may enhance consumer awareness of the artist. A larger stock of informed consumers could increase demand for her future albums. On the other hand, if consumers have a taste for diversity for albums by different artists, then the new release is likely to reduce demand for her future albums. The *backward spillover* refers to the impact of the release of a new album on sales of previously released albums by that artist. An album is classified as catalogue approximately 12-18 months after its release. If the new release increases the stock of informed consumers, then it could cause an increase in sales of catalogue albums. On the other hand, if consumers prefer new albums to old albums, the new release could cannibalize sales of catalogue albums.

Figure 1 illustrates how we can use the variation in the consumer's choice set over time to identify the backward spillover. 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 catalogue titles increase substantially. In the case of the "Bloodhound Gang,"a relatively obscure alternative rock band, the second album was considerably more popular than the first, and its release catapulted sales of the prior album to levels even higher than it had attained at the time of its own release. For the "Foo Fighters,"a more popular hard rock band with a very successful debut album, the impact of the second release is somewhat less dramatic, but still seems to have generated a substantial increase in sales of the band's first album. In both examples, the spillover effect appears to begin in the weeks just prior to the new album's release, and it persists for many months. In fact, for the "Bloodhound Gang," the effect persisted for at least three years.

Our main goals in this paper are to measure the backward spillover for artists and determine whether it is positive or negative, temporary or permanent. We study these questions using a large sample of recording artists whose debut albums were released in the United States during the period 1993 to 2002. Many of the artists in our sample released as many as four albums during the sample period, allowing us to study the variation in backward spillovers over different pairs of albums (e.g., release of albums 3 and 4 on sales of albums 1 and 2, and album 4 on sales of album 3). The variation in the data does not allow us to measure the forward spillovers. If artists had released their albums differently in separate geographical markets, then it may have been possible to identify the forward spillover between a pair of albums from the backward spillover between the same pair of albums in a market where their order of release is reversed. Unfortunately, in our sample, artists always released their new albums in the same order, and at the same time, in every geographical market.

The motivation for studying backward spillovers between albums is their impact on the relationship between the artist and distributor[1] and on market structure. The presence of a forward spillover implies that the distributor faces a standard holdup problem. If the debut album is a hit, and there is no contractual commitment, the artist can appropriate some of the surplus generated by the distributor's investment in the album by selling off the rights to future albums. Of course, this will reduce the distributor's incentives for investing in the debut album. (In practise, the typical contract gives the distributor the option to release future albums by the artist at the same terms, but terms are almost always renegotiated following a hit.) The backward spillover implies that the artist faces a lock-in problem. If she tries to sell the rights to the new release, the incumbent distributor (i.e., the one that financed her previous albums) has an advantage. It internalizes the impact of its investment in the new release on catalogue sales, whereas its rivals do not. Hence, the incumbent's willingness to invest in the new album, and to pay for it, exceeds that of its rivals. The incumbent can exploit this competitive advantage by extracting some of the surplus on the new release. Thus, even if artists and distributors cannot commit to long-term contracts, artists will rarely switch distributors, and this indeed appears to be the case in our sample. The magnitude of the backward spillover helps determine the relative bargaining power of the artist and the distributor, and therefore investment efficiency. It also affects entry. If a new distributor cannot easily bid away established artists, then the only way it can enter the industry is by developing its own stable of artists, which can take a long time, and may be impossible to accomplish without a prior reputation from having produced hit artists before.

Our empirical strategy for quantifying the backward spillover is taken from the literature on treatment effects. A new album release by an artist is interpreted as "treatment," and we are interested in measuring the difference between sales of catalogue titles for artists with

[1] royalty contract in which the artist usually gets 10% to 20% of album sales after the distributor recovers album allowances.

vs. without treatment. Since the artist cannot be in both states at the same time, we only observe one of the outcomes. However, if release times are random, then sales of catalogue titles for "untreated" artists with the same number of catalogue albums can be used to estimate the counterfactual sales for "treated" artists. Of course, the variation in release times across artists may not be entirely random. We use fixed effects to control for time-invariant factors such as genre and artist popularity that may influence release times, and conduct various checks to determine the robustness of our results. We find that the average treatment effects are permanent and substantial. The effect of the second album on sales of the debut album during the first six months of the treatment period is approximately 50% per month, and the effect shows no indication of declining with time. The treatment effects of the third and fourth albums on catalogue titles are smaller but significant, and they also appear to be permanent. They range between 20-30% in each of the first six months of the treatment period. We also explore heterogeneity in the magnitudes of the spillovers across artist characteristics.

Our analysis sheds light on the sources of the spillover. We show that the spillovers are not due to higher arrival rates of consumers who were previously not aware of the catalogue albums. The estimates of the short-run and long-run impact of the new release on catalogue sales, and the variation in these estimates across album pairs, are consistent with a model in which consumers buy catalogue albums because the new release increases their expected utility from these albums and hence their probability of purchasing the album. The increase in expected utility may reflect preference complementarities or new information. In the former case, the new album increase the utility of catalogue albums directly through consumption externalities and, in the latter case, it does so by generating signals that stochastically increases consumer beliefs about the value of purchasing the album. Our reduced form analysis does not provide a definitive answer to this question. However, it establishes a set of facts that should prove useful in formulating and estimating a structural model of demand and learning.

The paper is organized as follows. Section 2 discusses the data and reports descriptive statistics. Section 3 outlines the empirical strategy. The main estimation results are reported in Section 4. Section 5 examines the robustness of the estimates and the variation in spillovers across artists' characteristics. We interpret the results in Section 6. Section 7 concludes.

2 Data

Our data describe the album sales histories of 355 music artists whose debut albums were released 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, most notably *allmusic.com*, were also consulted for auxiliary information about genres and record labels and to verify album release dates.

The sample was constructed by first identifying a set of candidate artists who appeared on Billboard's "Heatseekers" chart, which lists the sales ranking of the top 25 new or ascendant artists each week.[2] 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 Heatseeker list was followed by a rise to stardom, we note (and show in detail below) that it also includes many "unknown" artists whose success was modest and/or fleeting. (The weekly sales of the lowest-ranked artist on the Heatseekers chart is typically around 3,000, which is only a fraction of typical weekly sales for releases by famous artists who have "graduated" from the Heatseekers category.)

Because our primary objective is to study demand responses to newly released albums, we restrict our attention to major studio releases. Singles, recordings of live performances, interviews, holiday albums, and anthologies or "greatest hits" albums are excluded from the analysis because they rarely contain any new music that could be expected to affect demand for previous albums.[3] The resulting sets of albums were compared against online sources of artist discographies to verify that we had sales data for each artists' complete album history; we dropped any artists for whom albums were missing or for which the sales

[2] 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.

[3] 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.

data were incomplete.[4] 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.[5] Finally, we narrowed the sample to artists for whom we observe the first 52 weeks of sales for at least the first two albums; we then include artists' third and fourth albums in the analysis if we observe at least the first 52 weeks of sales for those albums (i.e., we include third and fourth albums if they were released before 2002).

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

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

Table 1 summarizes various important patterns in 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

[4] 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.

[5] 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.

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. Recording and distribution costs for a typical album are in the ballpark of $200,000-$300,000, so an album must sell roughly 15,000 units (at around $16 per unit) in order to be barely profitable; most of the albums in our sample passed that threshold in the first year. However, although most of the albums in the sample were nominally successful, the distribution of success is highly skewed: as the table illustrates, sales of the most popular albums are orders of magnitude higher than sales of the least popular ones. For debut albums, for example, first-year sales at the 90^{th} percentile first release are ten times sales at the median and over 100 times sales of the album at the 10^{th} percentile.

The skewness of returns is even greater across artists than across albums, since artist popularity tends to be somewhat persistent. An artist whose debut album is a hit is likely to also have a highly successful second release, so that 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 for 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.

Figure reffig:typicalpaths shows "typical" sales paths for first and second releases, as well as the typical timing of later releases. (The paths depicted are kernel regressions of monthly sales on time since release; the vertical lines marking subsequent release dates reflect median time to release.) Although there is obviously heterogeneity across albums—not every album's sales path looks like the one in the picture—the figure conveys the predominant pattern: an early peak followed by a steady, roughly exponential decline. As indicated in the third and fourth panels of table, sales typically peak in the very first week and are heavily "front-loaded": a large fraction of the total sales occur in the first four weeks after release. Debut albums are an exception: first releases sometimes peak after several weeks, which presumably reflects a more gradual diffusion of information about albums by new artists. The degree to which sales are front-loaded seems to increase with each successive release.

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

The last panel of table reftable:summary 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 3 shows a more complete picture of the heterogeneity in release lags across albums, including elapsed time between non-adjacent albums. The distribution of elapsed time between albums 1 and 2 is clearly very similar to the distribution between albums 2 and 3, but the right truncation is obvious in cases involving the release of album 4.

Variation in the time it takes to release a new album could have several sources. Presumably there is a great deal of randomness in the creative process: developing new music requires ideas, coordination, and effort, all of which are subject to the vagaries of the artist's moods and incentives. Live tours and other engagements may delay the production of an artist's next album, and in some cases artists may "stall" for strategic reasons. (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' only leverage for negotiating more favorable terms in these contracts is to threaten to withhold new music.) Even when artists cooperate, it is possible that their record companies try to strategically time releases of new albums. Anecdotally, some record company executives talk of timing releases so as not to cannibalize sales of previous albums, and the seasonality in album release dates shown in Table reftable:releasemonths suggests that at least some discretion is exercised by the record companies.

Our empirical strategy involves comparing the sales of albums whose artists have recently released another new album to the sales of albums whose artists have not yet released another album, so it is critical that the timing of new release be exogenous with respect to sales dynamics. Although some of the anecdotal evidence mentioned above suggests that release times are endogenous, whether any such dependence is important is an empirical question. Appendix 1 summarizes an investigation of observable factors that could possibly influence time to release. The time it takes to release a new album appears to be independent of the success of the current album, which is somewhat surprising given that popular albums are more likely to be followed by live tours. Country artists typically release new albums sooner than rock or rap artists, and the time trend over the sample period appears to be toward shorter delays between releases. Of principal relevance to our empirical analysis is the relationship between time to release and the "shape" of an album's sales path. We show in the appendix that albums exhibiting faster decline rates are associated with longer delays before the release of the subsequent album. The implications of this relationship for our estimates of demand spillovers are discussed in section 4.

3 Empirical Strategy

Our empirical strategy is taken from the literature on treatment effects.[6] A new album release by an artist is interpreted as "treatment". Releasing a new album is an irreversible act: once treated, the catalogue albums remain treated. We will follow the impact of a new release on sales of catalogue titles for S periods, and refer to this number as the length of the treatment "window". Each new release is analyzed as a separate treatment episode. For each episode, time is measured in terms of the number of periods since the last new

[6] See Wooldridge [4]

album was released.

Without loss of generality, we focus on the first treatment episode. Let y_{it}^0 denote log of album 1 sales of artist i in period t without treatment and let y_{it}^s denote log of album 1 sales in period t when artist i is in the s^{th} period of treatment. Our objective is to estimate the average treatment effect (ATE) for each period of the treatment window:

$$ATE_s = E[y_i^s - y_i^0], \; s = 1, .., S.$$

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 nonlinear: a new release has a larger impact on total sales of catalogue 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.

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

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

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

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

Here μ^s is the mean of the distribution of log sales in time period t for artists in the s^{th} period of treatment, $\phi(t)$ is a function that that captures the common, downward trend in an artist sales, ν_i measures the impact of unobserved artist characteristics on sales in every period, and v_{it}^s is the idiosyncratic shock to album 1 sales of artist i when she is in treatment period s at time period t. The artist-specific effect does not vary across the treatment window. Substituting the above equations, the observed outcome for artist i in period t is given by

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

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

We use the outcomes of a random sample of artists as proxies for the missing sales data on each artist. 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.[7]

The regression model is as follows:

$$y_{it} = \alpha_0 + \alpha_i + \theta_t + \sum_{m=2}^{12} \delta_m D_{it}^m + \sum_{s=-3}^{5} \beta_s I_{it}^s + \epsilon_{it}, \tag{1}$$

where α_i is an artist fixed effect, the θ_t's are time dummies, and the D^m's are month-of-year dummies (to control for seasonality).[8] Here I_{it}^s is an indicator equal to one if the release of artist i's new album was s months away from period t, so β_s measures the new album's sales impact in month s of the treatment window. Intuitively, after accounting for time and artist fixed effects, we compute the difference in the average sales of album

[7] 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.

[8] Because we are using data at the 4-weekly frequency, time periods may span two calendar months, so the month-of-year indicators are not zeros and ones. Instead, we calculate the fraction of the time period associated with each month: e.g., if time period t for artist i's album included the last week of November and the first three weeks of December, then $D_{it}^{11} = 0.25$ and $D_{it}^{12} = 0.75$.

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 (θ_t) allow for a flexible decay path of sales, but implicitly we are assuming that the shape of this decay path is the same across albums: although differences in the level of demand are absorbed in the album fixed effects, differences in the shapes of albums' time paths are necessarily part of the error (ϵ). Including separate indicators for successive months 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 spillovers reflect intertemporal demand shifts. We allow for a 9-month treatment window, beginning three 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 demand spillovers reflect consumers' responses to these pre-release marketing campaigns. Second, including pre-release dummies serves as a reality check: we consider it rather implausible that a new album could have an impact on prior albums' sales many months in advance of its actual release, so if the estimated "effects" of the pre-release dummies are statistical zeros for months far enough back, we can interpret this as an indirect validation of our empirical model.

The regressions yield consistent, unbiased estimates of the treatment effect if the treatment indicators 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, treatment is random across artists. This is a strong assumption but not implausible. Selection effects would arise if the distributor uses a release rule that depends upon the flow of sales. For example, suppose the rule is to release the second album if sales of album 1 reaches a certain (random) threshold. In this case, the probability of treatment is more likely following a bad sales shock. However, it is not clear that such rules make sense from a decision-theoretic perspective. If the new release does not cannibalize sales of catalogue titles, then the distributor (and artist) should try to release a new album as soon as possible. We suspect that the main factor determining the time between releases is the creative process, which is arguably exogenous to time-varying factors.

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

4 Results

Table 3 presents GLS estimates of equation 1, obtained under the assumption of 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 nine months of the treatment window (i.e., the $\hat{\beta}_s$'s). Since the dependent variable is the logarithm of sales, the coefficients can be interpreted as approximate percentage changes in sales resulting from the new release.

In each treatment episode, the estimated impact of the new album three months prior to its actual release is statistically indistinguishable from zero. As discussed above, this provides some reassurance about the model's assumptions: three months prior to the treatment, the sales of 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 two months prior to the new album's release, growing in magnitude until the month of the release ($t = 0$ in the table), at which point there is a substantial spike upward in sales.

The estimates of the spillovers for each of the five months following the release of a new album are always positive, substantive, and statistically significant. In fact, for most album pairs (and especially the impact of album 2 on album 1) the estimated coefficients indicate a remarkable degree of persistence: the spillovers do not appear to be transitory. The only apparent exception is the impact of album 3 on album 2, for which the coefficients decline somewhat at the end of the treatment window. It is important to note, however, that the increasing coefficients in some specifications do not imply ever-increasing sales paths, since the treatment effects in general do not dominate the underlying month-to-month decay 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 largest spillovers are between albums 2 and 1, with estimates ranging between 40-55%. The spillovers for the remaining pairs of albums are substantially smaller, ranging mostly between 20-30%. The magnitudes are remarkably similar. There is some evidence that the spillovers from albums 3 and 4 onto earlier albums decline with each successive album: e.g., the impact of album 3 on the sales of album 1 is larger than the impact of album 4 on album 1, but the decline is not uniform and marginal at best.

The economic significance of the estimates of proportional changes reported in Table 3 are not immediately obvious since sales of the albums decline steadily with time. Table 4 shows the implied increases in total sales over the 9-month treatment window for catalogue albums for each new release. The increases are reported at the 10^{th}, 50^{th}, and 90^{th} percentiles of the sales distribution at the time of the respective new release. (In each case we compute numbers at the median release period.) We report these percentiles because the level of sales across albums is extremely heterogeneous, so the proportional effects listed in table 3 imply quite different increases in absolute sales for different artists. During the sample period, the retail price of albums is typically between $10-$18.

The implied effects are very large for hit albums—for example, a 90^{th} percentile first album sells over 40,000 extra units in the months following the release of the second album—and the numbers are economically meaningful: multiplied by a ballpark CD price of $16, the estimated effects imply revenue increases of over $700,000. On the other hand, the implied sales increases are inconsequential for the very small albums: even a large proportional increase doesn't mean much when applied to relatively low sales flows. The numbers in the table make clear that the biggest spillovers are between adjacent albums—e.g., albums 2 and 1, 3 and 2, or 4 and 3. Also, the implied effects are still large at the release of the fourth album.

4.1 Robustness

1. Robustness with respect to randomness of treatment: report the results of the regressions of differences in sales.

2. Robustness with respect to the proportional specification: split the sample into above median and below median albums and report the results.

3. Sample selection issues

4.2 Spillovers and Artist Heterogeneity

1. Early treatment versus late treatment.

2. Genre effects

3. Treatment on Hits versus Duds

4. Treatment from Hits versus Duds

5 Interpretation

The results suggest that the release of a new album shifts the sales path of a catalogue album up by a constant percentage, at least over the five months following the release. Since sales of a catalogue album decline with time, the spillover of a new release on catalogue albums is higher for later albums. However, we need to interpret these results with some caution. One implication would appear to be that releasing an album much earlier could increase total sales of catalogue titles. For example, if album 2 were released at six months rather than 18 months, the spillover would be much larger because it would be applied to a much higher sales level for album 1. It is important, however, not to extrapolate the results beyond the sample. In the data, times between releases are almost always at least a year and often longer. The intervals are sufficiently long that sales of most catalogue titles are fairly flat when the new albums are released. This implies that varying release times within the time frame observed in the data is unlikely to affect total sales.

The magnitude of the spillovers of later albums, such as the spillover of the fourth album onto the third album, and the variation across pairs of albums, argues against an artist discovery process being the main determinant of the spillovers. Consumers who would ever care to purchase the music of an artist will typically be aware of that artist by the time of the fourth release[9]. It is difficult to believe that the fourth release can cause tens of thousands of consumers to hear about a popular artist for the first time. Furthermore, even if this were true, it would seem to imply that the spillover of the fourth album onto sales of the first and second albums would be similar in size to its spillover onto the third album. However, the latter spillover is typically 2-3 times larger than the spillovers onto the first and second albums.

The variation in spillovers across pairs of albums measured in both levels and percentages is consistent with a model of album discovery. In the appendix, we develop a model in which consumer learn about the existence of an album over time, and the new release increases their arrival rates. However, this model fails to explain why the spillover measured in percentage terms is constant over the treatment window. As we show in the appendix, arrival rates over the treatment window would have to increase at an increasing rate in order to offset the more rapid decline in the stocks of potential consumers. This hypothesis seems implausible, and implies ever smaller spillovers for subsequent releases. If changes in arrival rates were the main source of the spillovers, we would have expected estimates of the treatment effect to decline over the treatment window. We find no evidence of a decline, even when we extend the treatment window to nine months.

The results become considerably easier to interpret if we assume that the new release increases the probability that consumers purchase catalogue albums upon learning about the album. This is true for two reasons. First, the increase in the probability applies to all consumers who have not yet bought the catalogue album, and not just to those who have not yet learned about the album. This can help explain the magnitude and persistence of the spillovers. Second, the increase in the probability of purchase is likely to be a constant. This can help explain the proportional shift in the post-treatment path of log sales of the

[9] This is not to say that fans of rock music will be familiar with country artists; but rather that potential buyers of country music (i.e., country fans) will be aware of any country artist that is releasing a fourth album.

catalogue album. The question then arises what factors can explain why the new release increases the probability with which consumers purchase catalogue albums. One possibility is that albums are complements. The new release increases the utility of catalogue albums and, as a result, some consumers who previously did not think the catalogue albums were worth purchasing will change their minds and buy the catalogue albums. Indeed, we show in the appendix that complementarities can explain the variation in spillover estimates across pairs of albums, and across the treatment window for each pair of albums. However, an alternative explanation is uncertainty about preferences. Consumers learn about their preferences for an album over time by listening to the album's songs on the radio, by listening to songs on other albums by the artist, by talking to friends, by observing the success of the artist measured in terms of numbers of albums and their Billboard ranking. All of these events provide information that leads consumers to update their beliefs about the album and can cause the expected utility of an album to change. Discriminating between these two models is a challenging task.

6 Conclusion

We find that the backward spillover of a new release on sales of catalogue albums are positive, substantial, and permanent. The effect of the second album on sales of the debut album during the first six months of its release period is approximately 40-50% per month, and the percentages show no indication of declining with time. The effects of the third and fourth albums on catalogue albums are smaller but still significant, approximately 20-30% in each of the first six months of the release period. The magnitudes are large enough to affect the relationship between artists and distributors, and to act as a barrier to entry for potential entrants.

The key remaining issue is to identify the sources of the spillover. We have argued that informational stories in which the new release increases the arrival rate of consumers who want to buy the artist's albums do not appear to be consistent with the data. The potential explanations are preference complementarities or preference uncertainty. In both models, the consumer's utility from an album can change over time even though her preferences do not. The distinguishing feature of preference complementarity is that the catalogue album is purchased if and only if the consumer also purchases the new release. If the new release simply provides additional information, the consumer may be just as likely to purchase the catalogue album as the new release. We hope to provide a more definitive analysis on the sources of the spillovers in our next paper using a structural model that exploits the variation in sales across geographical markets as well as across artists.

7 References

References

[1] Becker, G., Michael Grossman, Kevin M. Murphy, “An Empirical Analysis of Cigarette Addiction”, American Economic Review, Vol. 84, No. 3, June, 1994, pp. 396-418.

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

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

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

8 Appendix

In this appendix, we present a model of album learning and album demand. The critical assumption of our model is that album awareness is a discrete event. When consumers become aware of the artist's new album, they know their preferences for the new album and for any albums released prior to it.

An artist produces K albums and releases them sequentially in a market with N consumers. Albums are indexed by k. The time period between releases are called release periods. For convenience, we will assume that release periods are the same length and equal to T. The first album is released at time 0 and album k is released at time $(k-1)T$. The release periods are labelled by the album released in that period. Album k is classified as catalogue in subsequent release periods. In order to distinguish between albums and release periods, we will use s to index the release periods. Thus, album k is released at the beginning of release period k but it can also be purchased in any later release period $s = k+1, .., K$ as a catalogue album. An album is a durable good that, once purchased, generates a flow of utility over time. Each album retails at a fixed price in every period following its release.

Let $d(s) = (d_1(s), ..., d_s(s))$ denote the consumption vector of albums consumed in period s, where $d_k(s)$ is equal to one if album k has been purchased prior to or in period s and zero otherwise. The consumer's preferences take the form

$$V(q_0) = \sum_{s=0}^{\infty} \delta^t [q_0(s) + U(d(s); \theta(s))],$$

where δ is the discount parameter $(0 < \delta < 1)$, $q_0(s)$ is consumption of the numeraire good in period s, U is a subutility function that measures the consumer's utility in period s from the consumer's stock of albums in period s, and $\theta(s) = (\theta_1, .., \theta_s)$ is a vector of idiosyncratic, album-specific preference shocks. In what follows, we will assume that

$$U(d(s)) = \sum_{k=1}^{s} d_k(s)\theta_k + \sum_{k=1}^{s} \sum_{l=1}^{k} \sigma_{kl}\theta_k\theta_l d_k(s) d_l(s).$$

We will refer to θ_k as the standalone flow utility that consumer derives from album k. The first term then is the sum of the standalone utilities of the albums purchased by period s; the second term captures possible interactions between the albums purchased. Preferences are submodular if σ_{kl} are negative for all k, l and supermodular if σ_{kl} are positive for all k, l. Let F denote the joint distribution of θ in the population with support $\Re_+^K$.

At the beginning of each release period s, every consumer gets an independent random draw from a distribution G_s that specifies the time at which she arrives in the market. Consumers who arrive before the end of the release period learn about the available set of albums and her preferences over the choice set. The instantaneous probability that a consumer arrives at time t in release period s is given by

$$h_s(t) = \frac{g_s(t)}{[1 - G_s(t)]}.$$

(Here t measures time since the latest album was released.) Note that if $G_s(T)$ is less than one, then there is a positive probability that a consumer may not learn about the new album released in that period.

We say that the consumer is *myopic* if she bases her purchasing decision in release period s solely in terms of her preferences over the choice set available in period s. She acts as if $U(d(s))$ is the flow utility that she will obtain in all subsequent periods, thereby ignoring the possibility that she may purchase subsequent albums and that these albums will affect the utility that she derives from $d(s)$. Let $\tau(\theta)$ denote the optimal purchasing rule for the myopic consumer with preferences θ and let $\tau_k(\theta)$ denote the release period in which the consumer θ buys album k. Then the probability that album k is purchased in release period s ($s \geq k$) can be defined by

$$\lambda_{k,s} = E\{I_{\{\tau_k(\theta)=s\}}\},$$

where I is the indicator function and the expectation is taken with respect to θ.

If U is additive, then the consumer's myopia is irrelevant since each album is judged solely in terms of its standalone utility. In that case, the optimal purchase plan is clearly to purchase album k in period k if and only if the standalone utility θ_k exceeds discounted price. It implies that the consumer purchases an album only in the period in which it is released. This result extends to the case of submodular utility functions.

Proposition 1 *Suppose U is submodular, consumers are myopic, and every consumer learns about an album in its release period. Then, for each $k = 1, .., K$, $\lambda_{k,s} > 0$ for $s = k$ and $\lambda_{k,s} = 0$ for $s > k$.*

The intuition is that subsequent album releases can only lower the utility of albums released previously, and hence increase the purchasing thresholds on θ that these albums must meet to be worthwhile purchasing. Of course, myopic consumers may regret prior purchases. Album k could yield more utility than an earlier release but, because the latter was purchased first, album k may not be worth purchasing. The forward-looking consumer will guard against these kinds of mistakes by delaying purchases. Thus, the purchasing thresholds for the rational consumer are likely to be higher than those of a myopic consumer. As δ increases, the cost of delaying purchases decreases, which causes purchasing thresholds to increase. Thus, the purchasing plans of myopic and rational consumer diverge as consumers becomes more patient. In fact, in the limit, the rational consumer simply waits until period T to make her purchasing decisions.

If U is supermodular, then future releases increases the flow of utility from current and past album purchases. In this case, the purchasing pattern for albums is quite different.

Proposition 2 *Suppose U is supermodular, consumers are myopic, and every consumer learns about an album in its release period. Then $\lambda_{k,s} > 0$ for $s = k, .., K$, and $k = 1, .., K$.*

The proposition states that each album is purchased with positive probability in every period after it is released. The intuition is that a new release always increases the incremental utility of albums released previously and hence decreases the θ-purchasing thresholds that these albums must satisfy to be worth purchasing. Note that this implies that the myopic consumer never regrets purchasing any album: as the choice set increases, he will still want to buy the albums that he had purchased earlier.[10] However, the timing of the purchases is not optimal. The forward-looking consumer whose utility from album k currently does not meet the threshold should anticipate the possibility that he may purchase later albums when they are released. Since these purchases will increase the utility that he derives from album k, he may find it worthwhile to purchase album k in later periods. If he does so, he will regret not purchasing album k earlier. The gain from waiting is more informed purchasing decisions but the cost is delayed consumption. The optimal plan will strive to equate the marginal gain to the marginal cost of waiting, which will lead consumers to purchase album k even if their utility from album k is currently slightly less than the price. In this case, however, as δ approaches 1, the cost of delayed consumption gets very small, and hence we conjecture that the purchasing plan of the rational consumer converges to that of the myopic consumer.

[10] One can show that the myopic consumer ends up making exactly the same purchases as he would have made had he waited until period K to make his decisions.

In practise, consumers may not learn about the new album in the period of its release. In fact, they may not learn about the album for several release periods. In that case, the set of potential consumers at time t in release period s is comprised of at most s cohorts. The cohorts are indexed by the release period of their latest arrival time. Let $N_{k,s}(t)$ denote the number of consumers at time t of release period s whose latest arrival time was in release period k, where $k < s$. The number of consumers who have not yet arrived by time t in release period s is given by $N_{0,s}(t)$. Thus, a member of the cohort (k, s) has had an opportunity to purchase albums 1 through k but consumer has not had an opportunity to purchase albums $k + 1$ through s.

Let $y_{k,s}(t)$ denote aggregate sales of album k at time t of release period s. In our empirical work, we measure sales frequencies at monthly intervals. Therefore, dividing each release period is divided into n periods of equal length, which we normalize to 1, we define the discrete analogue of the hazard rate by

$$\alpha_s(t) = \exp\left[\int_t^{t+1} h_s(u)du\right].$$

Proposition 3 *Suppose preferences are additive. Then the expected sales rate of album k in release period s $(s \geq k)$ is given by*

$$y_{k,s}(t) = \alpha_s(t)\lambda_{k,k}\sum_{r=0}^{k-1} N_{r,s}(t) \tag{2}$$

The formula for sales follow from the fact that arrival times are independent of preferences and the assumption that a sufficient number of consumers make purchasing decisions during period t that we can appeal to the law of large numbers to argue that the fraction of consumers who purchase the album is approximately equal to $\lambda_{k,k}$, the probability that the standalone utility of a randomly selected consumer for album k exceeds the discounted price. The risk set of potential buyers consists of consumers who have not had an opportunity to purchase album k, which is simply the number of consumers whose latest arrival time occurred before release period k.

Since the purchasing probabilities are fixed, a positive spillover occurs in this model only if new releases increases the arrival rate of consumers who are interested in buying the artist's albums. In that case, the variation in spillovers across catalogue albums is easily explained in terms of the variation in the risk set. The spillover of album s on catalogue album k is larger than its spillover on album $k-1$ since the risk set of the latter excludes the cohort of consumers whose latest arrival time was in release period k. However, the model fails to explain the constant percentage increase in sales of catalogue albums across the treatment window. The higher arrival rates caused by the new release implies a faster decline in the risk set of potential buyers. To offset this decline, the hazard rate has to increase over time at an increasing rate. This leads to a contradiction.

By contrast, suppose arrival rates of consumers are exogenous at rate α but preferences are supermodular. In this case, spillovers occur because the new album increases probability of purchase. Consumers who have not yet learned about the catalogue album are more likely to buy it when they do learn about it; consumers who know about the catalogue album but did not find it worth purchasing previously are now more likely to buy it when they can bundle it with the new album. Let $m_{k,s}$ denote the fraction of the population in cohort (r, s) when album s is released and let $R_{k,s}(t)$ denote the percentage increase in sales of album k in period t following the release of album s.

Proposition 4 *Suppose preferences are supermodular and $\alpha_s(t) = \alpha$. Then expected sales rate of album k at time t in period s is given by*

$$y_{k,s}(t) = \alpha \sum_{r=0}^{s-1} \left(\sum_{l=\max\{r,k\}}^{s} \lambda_{k,l} \right) N_{r,s}(t) \tag{3}$$

and

$$R_{r,s}(t) = \frac{\lambda_{k,s} \left(\sum_{r=0}^{s-1} m_{r,s} \right)}{\sum_{r=0}^{s-1} \left(\sum_{l=\max\{r,k\}}^{s} \lambda_{k,l} \right) m_{r,s}}$$

The formula for monthly sales in Proposition 3 differs from the sales equation in Proposition 2 in two important ways. First, the probability of purchasing album k is not a constant but varies across the cohorts. The summation inside the parentheses gives the probability that a consumer in cohort (r, s) purchases album k upon arrival in period s. It reflects the fact that other albums have been released since this consumer's latest arrival time in release period r. Second, the risk set of potential buyers consists of all cohorts, including those whose latest arrival time occurred *after* album k was released. The reason is that other albums have been released in the intervening periods, including album s, and these releases affect the consumer's probability of purchasing album k. For every cohort of potential buyers, the probability of purchase increases by $\lambda_{r,s}$.

The second part of Proposition 3 states that the percentage increase in sales of a catalogue album following the release of a new album is a constant percentage of what sales would have been had the new album not been released. The intuition behind this result is the increase in purchasing probabilities for each cohort is a constant and the size of the different cohorts are changing at the same rate. The spillover of album s on catalogue album k will be larger than its spillover on album $k-1$ for the same reason as in the previous model: the risk set of the latter excludes the cohort of consumers whose latest arrival time was in release period k. The spillovers measured in percentage terms depends primarily on the distribution of preferences. For example, suppose $s=3$ and assume that most consumer have learned about album 1 when the artist releases her third album. Then $R_{1,3} \lesseqgtr R_{2,3}$ depending upon whether

$$\frac{\lambda_{1,3}}{\lambda_{1,2}} \lesseqgtr \frac{\lambda_{2,3}}{\lambda_{2,2}}.$$

The ratios are both less than one, but it is not clear which is larger. Of course, even if the percentages are similar, the spillover on album 2 measured in levels is larger than the spillover on album 1 since the size of the risk set for album 1 is significantly smaller than that of album 2.

The conclusion that we draw from this exercise is that a model which attributes the spillovers to increases in purchasing probabilities fits the data much better than a model which attributes the spillovers to increases in arrival rates.

Figure 1: Typical sales paths

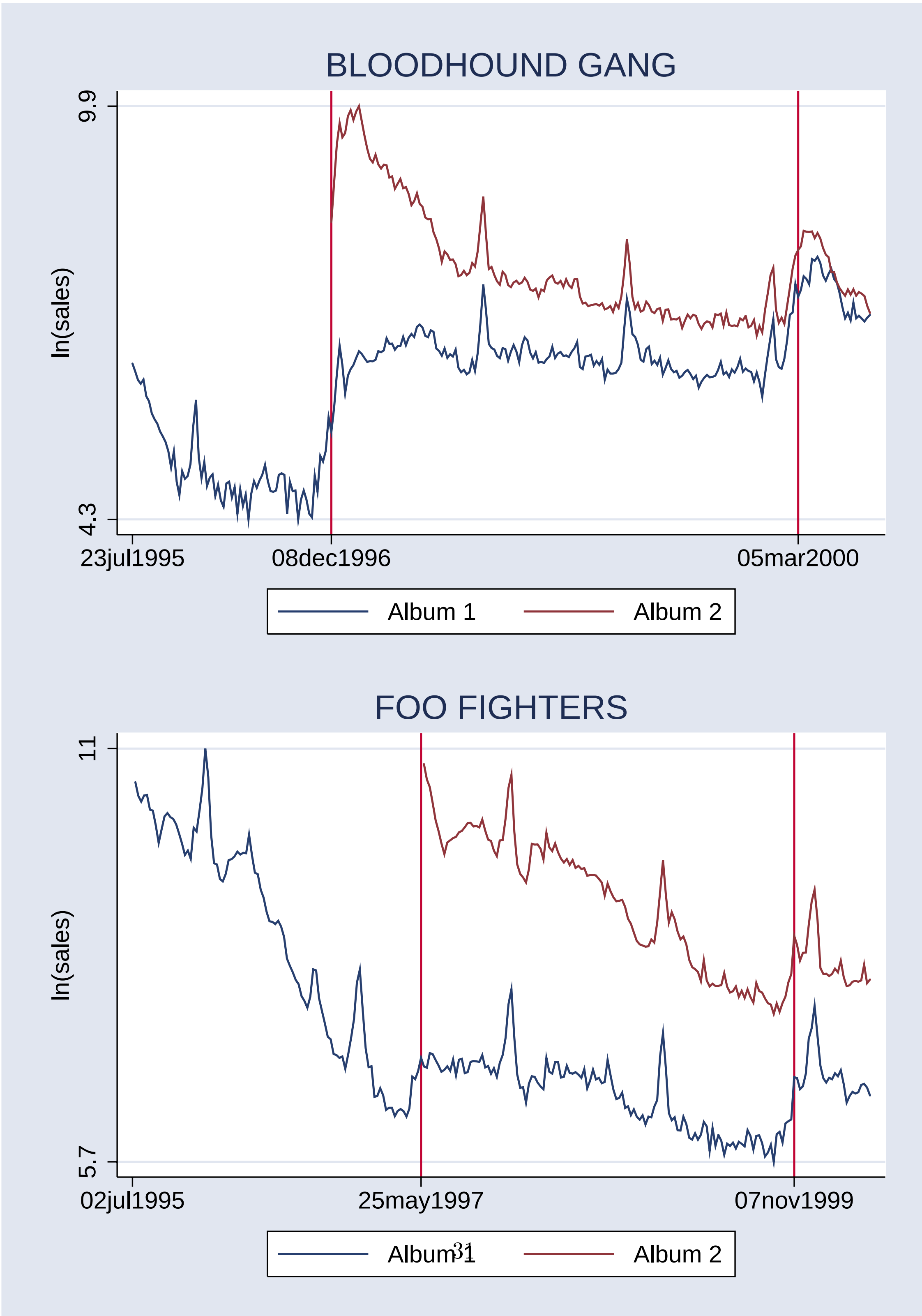

Figure 2: Typical sales paths

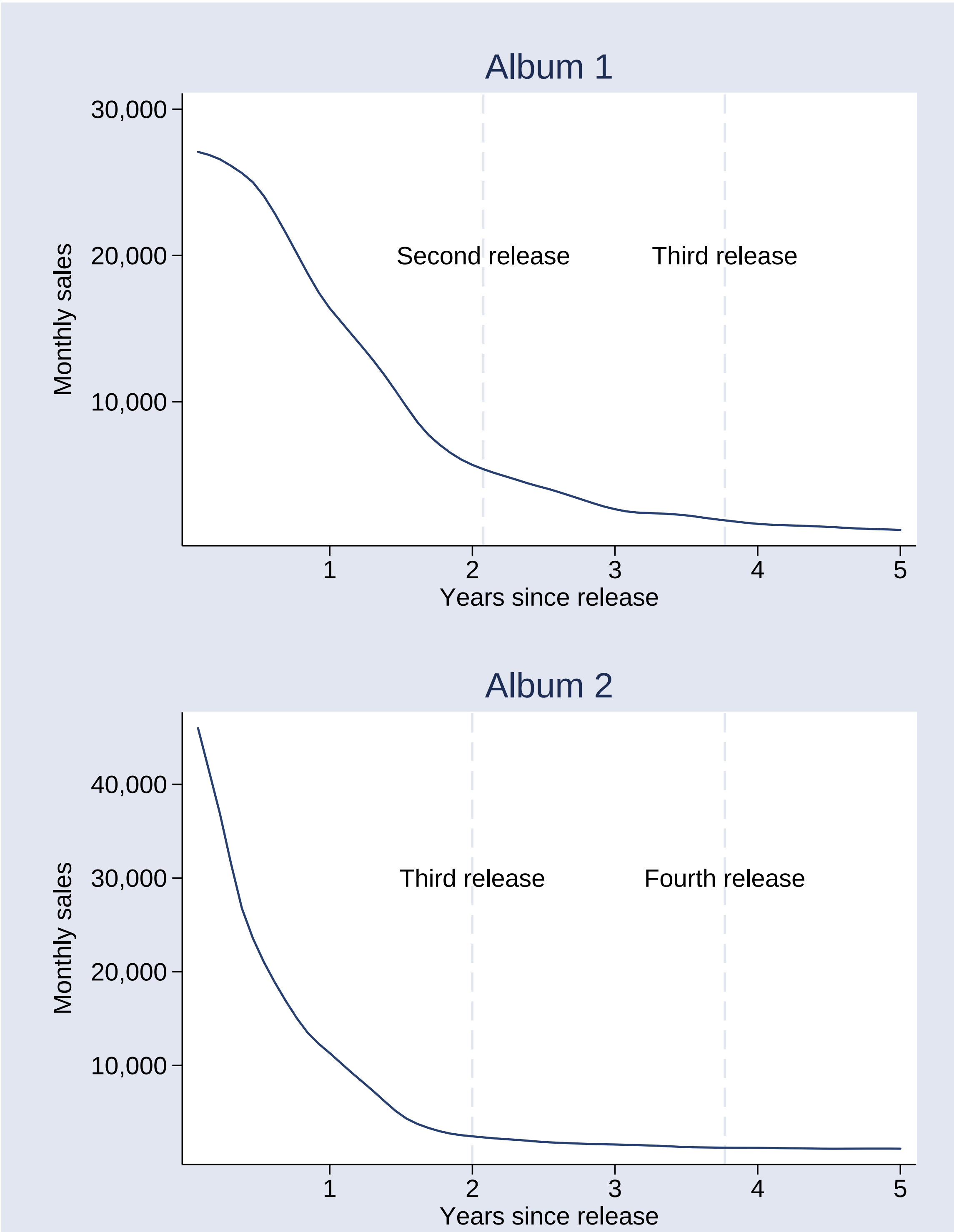

Figure 3: Distributions of Elapsed Time Between Releases

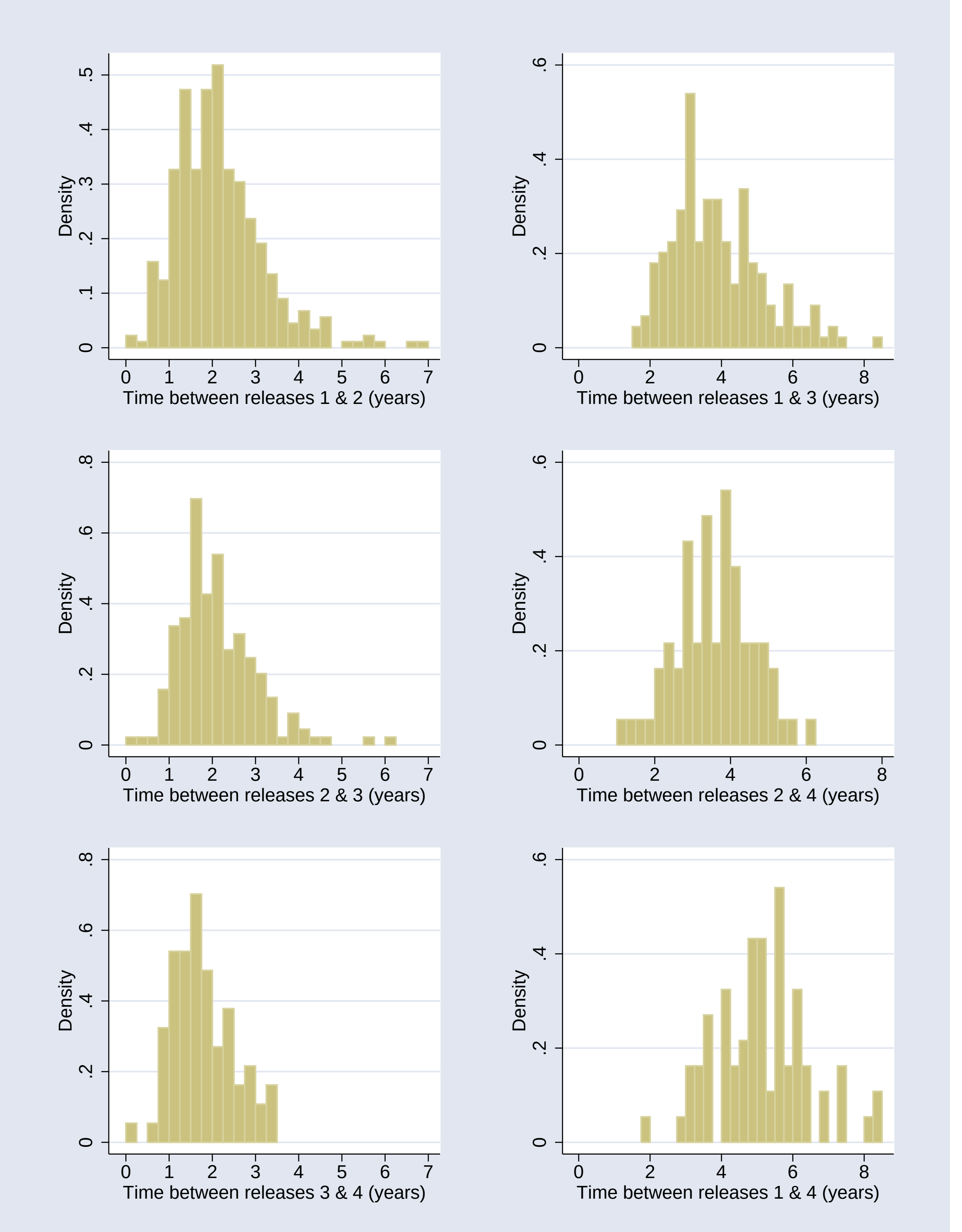

Table 1: Summary Statistics

	N	Mean	Std. Dev.	Percentiles .10	Percentiles .50	Percentiles .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				
	Album 1	Album 2	Album 3	Album 4	Overall
Month	(n=355)	(n=355)	(n=178)	(n=74)	(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: Estimated Effects of New Releases on Sales of Catalog Albums

Month (relative to release date)	Album pair					
	2→1	3→1	4→1	3→2	4→2	4→3
$t = -3$	.019	.017	.039	- .008	.009	.030
	(.015)	(.017)	(.026)	(.018)	(.023)	(.027)
$t = -2$	.082	.059	.081	.043	.058	.101
	(.019)	(.021)	(.031)	(.024)	(.028)	(.034)
$t = -1$	.235	.183	.180	.187	.167	.188
	(.022)	(.024)	(.034)	(.028)	(.030)	(.038)
$t = 0$	.428	.327	.292	.305	.283	.283
	(.024)	(.026)	(.035)	(.031)	(.031)	(.041)
$t = 1$	.467	.324	.294	.284	.252	.243
	(.026)	(.027)	(.037)	(.033)	(.032)	(.043)
$t = 2$	.497	.288	.311	.236	.273	.245
	(.027)	(.028)	(.038)	(.034)	(.033)	(.045)
$t = 3$	.502	.304	.281	.226	.204	.217
	(.029)	(.029)	(.039)	(.036)	(.034)	(.047)
$t = 4$	.503	.235	.320	.179	.230	.274
	(.029)	(.030)	(.040)	(.037)	(.035)	(.048)
$t = 5$	.533	.254	.325	.144	.230	.232
	(.030)	(.031)	(.041)	(.039)	(.036)	(.050)
# albums	338	162	66	173	70	74
# observations	9,027	4,461	1,582	4,585	1,664	1,696
$\hat{\rho}$	.749	.656	.524	.728	.505	.610

GLS estimates of the regression described in equation 1; standard errors in parentheses. Estimated coefficients for time dummies and seasonal dummies are suppressed to save space. Each column of the table represents an album pair: e.g., the column labeled 4→2 lists the estimated effects of album 4's release on the sales of album 2. The $t = 0$ month is the first month following the release of the new

Table 4: Implied Total Increase in Sales

	Level of sales prior to new release (percentile)		
Album pair	0.10	0.50	0.90
2→1	829	4,425	46,807
3→1	149	1,287	11,015
4→1	64	712	8,021
3→2	366	1,953	20,653
4→2	125	1,083	9,270
4→3	438	2,648	37,898

Total increase in sales over the 9-period treatment window implied by the estimates in Table 3. So, for example, the release of the third album increases sales of the second album by 1,953 units if the second album's sales were at median levels prior to the new release.

Spillover Estimates: First-differenced model

Month (relative to release date)	Album pair					
	2→1	3→1	4→1	3→2	4→2	4→3
$t = -3$	-0.003	0.026	0.048	-0.001	0.026	0.041
	(0.014)	(0.016)	(0.026)	(0.018)	(0.023)	(0.027)
$t = -2$	0.046	0.045	0.042	0.050	0.034	0.069
	(0.014)	(0.016)	(0.025)	(0.018)	(0.023)	(0.026)
$t = -1$	0.130	0.132	0.107	0.156	0.120	0.095
	(0.014)	(0.017)	(0.026)	(0.018)	(0.024)	(0.027)
$t = 0$	0.172	0.153	0.121	0.133	0.120	0.114
	(0.014)	(0.017)	(0.026)	(0.018)	(0.024)	(0.027)
$t = 1$	0.019	-0.003	0.000	-0.015	-0.025	-0.037
	(0.014)	(0.017)	(0.026)	(0.018)	(0.024)	(0.027)
$t = 2$	0.008	-0.030	0.017	-0.034	0.038	0.012
	(0.014)	(0.017)	(0.026)	(0.018)	(0.024)	(0.027)
$t = 3$	-0.018	0.014	-0.038	-0.014	-0.073	-0.021
	(0.014)	(0.017)	(0.027)	(0.018)	(0.024)	(0.028)
$t = 4$	0.003	-0.081	0.049	-0.039	0.036	0.061
	(0.014)	(0.017)	(0.027)	(0.019)	(0.024)	(0.028)
$t = 5$	0.009	0.011	0.009	-0.021	-0.002	-0.026
	(0.014)	(0.017)	(0.028)	(0.019)	(0.025)	(0.029)
# albums	338	162	66	173	70	74
# observations	9,010	4,451	1,579	4,577	1,661	1,695

Spillovers on Big vs. Small Albums: Album 1

Month (relative to release date)	2→1 Big	2→1 Small	3→1 Big	3→1 Small	4→1 Big	4→1 Small
$t = -3$	-.020	.032	.027	.002	.029	.072
	(.018)	(.022)	(.020)	(.029)	(.032)	(.042)
$t = -2$	.006	.131	.085	.023	.053	.142
	(.024)	(.029)	(.026)	(.036)	(.039)	(.050)
$t = -1$	.117	.305	.188	.169	.123	.296
	(.028)	(.033)	(.030)	(.041)	(.042)	(.053)
$t = 0$	.315	.461	.338	.297	.240	.398
	(.031)	(.036)	(.032)	(.043)	(.045)	(.055)
$t = 1$	.328	.514	.347	.270	.230	.425
	(.033)	(.038)	(.034)	(.045)	(.046)	(.057)
$t = 2$	.327	.563	.301	.280	.289	.422
	(.035)	(.040)	(.035)	(.047)	(.048)	(.058)
$t = 3$	.313	.567	.340	.266	.268	.356
	(.037)	(.041)	(.037)	(.049)	(.050)	(.060)
$t = 4$	.309	.572	.289	.149	.302	.334
	(.038)	(.043)	(.038)	(.051)	(.051)	(.061)
$t = 5$	.308	.609	.292	.193	.242	.401
	(.039)	(.044)	(.039)	(.052)	(.053)	(.062)
# albums	175	163	82	80	31	35
# observations	5,067	3,960	2,319	2,142	751	831

GLS estimates of the regression described in equation **??**; standard errors in parentheses. Estimated coefficients for time dummies and seasonal dummies are suppressed to save space. The "big" vs. "small" distinction is based on first-year sales: debut albums with first-year sales above the median (for debut albums) are big, and below-median albums are small.

Spillovers on Big vs. Small Albums: Albums 2 and 3

Month (relative to release date)	3→2 Big	3→2 Small	4→2 Big	4→2 Small	4→3 Big	4→3 Small
$t = -3$	-.031	.038	-.024	.071	.003	-.008
	(.022)	(.032)	(.026)	(.045)	(.034)	(.043)
$t = -2$	-.003	.118	.032	.077	.094	.006
	(.029)	(.040)	(.032)	(.051)	(.045)	(.050)
$t = -1$	.140	.254	.146	.156	.184	.083
	(.034)	(.043)	(.035)	(.054)	(.050)	(.054)
$t = 0$	.245	.389	.256	.284	.237	.219
	(.038)	(.046)	(.036)	(.056)	(.054)	(.056)
$t = 1$	.196	.416	.202	.335	.167	.212
	(.041)	(.047)	(.038)	(.057)	(.058)	(.059)
$t = 2$	.150	.368	.260	.292	.169	.207
	(.043)	(.049)	(.039)	(.057)	(.061)	(.061)
$t = 3$	.147	.332	.180	.252	.121	.155
	(.046)	(.051)	(.040)	(.059)	(.063)	(.061)
$t = 4$	.076	.332	.192	.326	.192	.217
	(.048)	(.052)	(.041)	(.060)	(.066)	(.063)
$t = 5$	.054	.284	.170	.267	.129	.164
	(.050)	(.054)	(.042)	(.063)	(.068)	(.065)
# albums	107	66	44	26	41	33
# observations	2,960	1,625	1,111	553	995	701

GLS estimates of the regression described in equation **??**; standard errors in parentheses. Estimated coefficients for time dummies and seasonal dummies are suppressed to save space. The "big" vs. "small" distinction is based on first-year sales: albums with first-year sales above the median are classified as big, and below-median albums are small.

Sample selection and spillovers:
4-release artists vs. 2- and 3-release artists

Month (relative to release date)	4-release artists	2- or 3-release artists
$t = -3$	-.018	.017
	(.032)	(.017)
$t = -2$	-.035	.086
	(.043)	(.022)
$t = -1$	.067	.245
	(.050)	(.025)
$t = 0$	.203	.446
	(.055)	(.027)
$t = 1$	.198	.490
	(.060)	(.029)
$t = 2$	.256	.503
	(.064)	(.031)
$t = 3$	.286	.495
	(.068)	(.032)
$t = 4$	.285	.489
	(.071)	(.033)
$t = 5$	.274	.521
	(.074)	(.034)
# observations	1,335	7,692

GLS estimates of the impact of album 2's release on album 1's sales; standard errors in parentheses. Estimated coefficients for time dummies and seasonal dummies are suppressed to save space..