Consumer Information in the Digital Age: Empirical Evidence from Spillovers in the Music Industry *

Sisley Maillard †

PRELIMINARY
April 2013

Abstract

In markets with abundant supply of goods, like books, music or movies, consumer choices depend not only on their preferences but also on their knowledge of the product space. With the development of online recommendation tools – online communities, social networks, personalization and recommendation technologies – it is expected that consumers will be better informed about the products that fit their tastes. In this paper, we conduct an empirical test of this hypothesis for the music industry. We measure consumer information through *backward spillovers*, which are the impact of a second album's release on the first album sales by the same artist (Hendricks and Sorensen, 2009). Since backward spillovers reflect consumers' lack of information about artists at the time of their first release, we study how the development of online recommendation tools affects the magnitude of backward spillovers. We use a dataset of weekly album sales in France for the period 2003-2010, and compare spillovers between a first sample of artists who released their debut album in the early stage of the digital age, in 2003, and a second sample of artists who debuted when online recommendation tools had become more widely available, in 2007. We find that, as compared with 2003, information spillovers have decreased in 2007, in a second stage of information dissemination. Our results imply that online recommendation tools have increased consumers consumer information via electronic word-of-mouth.

Keywords: Information; Spillovers; Digitization; Music industry; Recommendation; Word-of-mouth; Cultural markets

JEL Classification: C23; D83; L15; L82; O30

---

† Department of Economics and Social Sciences, Telecom ParisTech - 46 Rue Barrault, 75013 Paris, France. Contact: sisley.maillard@telecom-paristech.fr

* *I am grateful for the advice of Professor Marc Bourreau and seminar participants at Telecom ParisTech (Paris, France). I have also benefited from comments and discussions with seminar and conference participants at the 2012 ACEI Conference (Kyoto, Japan) and the 4th ICT Workshop (Évora, Portugal).*

1. Introduction

In markets, like books, movies, or music, consumers are exposed to an extensive supply of goods, with many new products entering the market each week.[1] Therefore, market demand depends not only on consumers' preferences but also on their knowledge of available products, and the process by which they obtain this knowledge.

This process is traditionally conducted through the mass media (i.e., radio plays and videos clips on television, as regards music albums). Consumers are often poorly informed about most products, and especially the new ones, since only a small set of products achieve visibility in the mass media – both because of capacity constraints and promotion bottleneck of the music industry. For example in France, about 3% of the music played on radio obtains 75% of the total radio plays.[2]

However, the new technologies of information and communication have been changing how consumers obtain information. According to consumer surveys, the Internet has been narrowing the gap with radio as the main channel for learning about new music,[3] and advices from acquaintances or opinions posted by consumers online are became the most trusted form of recommendations.[4]

In this paper, we study how digital technologies have been impacting product discovery, in the demand for recorded music. We assess empirically whether consumers are better informed about products that fit their tastes as online recommendation tools – including online communities, social networks, or personalization and recommendation technologies[5] – become more widely available.

Measuring how consumer information evolves is important in the extent that the main factor determining market demand is whether consumers know about the products. As shown by Hendricks and Sorensen (2009) in the pre-digital US music market, incomplete information about the choice set of albums contributes to the "skewness of music sales" (i.e.,

[1] More than 62,000 new albums were launched in France in 2010 (*Observatoire de la Musique/GfK*, 2010)
[2] *Observatoire de la musique* (from 2003 to 2010).
[3] In 2002, respondents cited radio (53%) as the first media to discover new music, followed by television (14%), the Internet (9%), and newspapers (2%). In 2010, they cited radio (39%) and Internet (31%), before television (12%) and newspaper (3%) as the main channel to discover new music (*Edison Research*, "National US surveys 2002-2010", 2010).
[4] Survey of 25,000 Internet consumers from 50 countries (*Global Advertising*, 2009).
[5] "Personalization and recommendation technologies take various forms, from simple top lists of recommended items other users have selected to sophisticated collaborative filters that infer preferences with little or no active effort by the consumer" (Brynjofsson et al., 2010).

a small number of very successful albums claims a large share of total music sales)[6] since consumers are more likely to be aware of popular products than niche products. Anecdotal evidence suggests that the skewness in album sales reflects the skewness in radio plays; for example, on 3,000 albums released in France in 2001, 30 were on radio playlists and 10 accounted for 80% of the total sales.[7]

There is an ongoing debate[8] about to what extent the development of online recommendation tools has been impacting consumer information, and thus market demand (Brynjolfsson et al., 2010). On the first hand, it could be argued that online recommendation tools allow consumers to acquire product information with greater convenience and at lower costs, leading to increased demand for less popular products which do not benefit from enough visibility (see e.g. Anderson, 2006; Brynjolfsson et al., 2006). On the other hand, there is some evidence that online recommendation tools further increase the informational inequality between popular and niche products, leading to reinforce consumer interest in the most popular products (see e.g. Fleder and Hosanagar, 2009; Dellarocas et al., 2010).

Our empirical strategy, to assess how consumer information has been evolving with online recommendation tools, is based on the learning-based model of market demand of Hendricks and Sorensen (2009). In their model, any promotional activity associated with a newly released album enhances consumer awareness about the artist and may cause some consumers to discover and purchase the artist's past albums, leading to an increase in sales of these albums. This effect, called the "backward spillovers", reflect consumers correcting initial mistakes and buying the first album at the time of the second release. We use backward spillovers to evaluate consumers' lack of information about their choice sets.

We use a data set of weekly album sales in France for the period 2003-2010, in order to compare the magnitude of backward spillovers at two stages of the digital development: between a first sample of artists who released their debut album in the early stage, in 2003, and a second sample of artists who debuted when online recommendation tools had become more widely available, in 2007. Indeed, social networks and online communities providing music recommendations have grown fast over the last decade, especially between 2003 and 2007,[9] to reach millions of users today.[10]

---

[6] This common pattern in cultural markets is also described by the classic Pareto Principle (20% of the products claim 80% of the total sales) or the theory of *SuperStars* (Rosen, 1981; Adler, 1985).

[7] *Epok*, n° 23 (edited by Fnac, a French retail chain specialized in cultural products).

[8] See for further details "Literature Review", section 2, p.4, §3.

[9] Including the music blogs publishers Skyblog (2002), the first music social network MySpace (2003), the video-sharing websites YouTube (2005) and Dailymotion (2005), the music blog aggregator Hype Machine (2005), the social network Facebook (2006), the micro-blogging platform Twitter (2006), the web radio and

The rest of the paper is organized as follows. Section 2 provides a summary of the related literature. Section 3 describes our data and Section 4 our empirical strategy. Section 5 presents our empirical results and discusses our findings. Section 6 gives some concluding remarks.

2. Literature Review

This paper is related to the literature on the impact of information provision on market outcomes. In markets with a large number of products whose quality is difficult to determine ex-ante, consumers face a problem of incomplete information about their choice set. Goeree (2008) has estimated a structural model of demand in the market for personal computer and shows that the rapid pace of technological change makes consumers less than fully informed about the set of available products. In a theoretical model, Cabral (2000) has studied the firms' decision to release new products under existing brand names when consumers are uncertain about product qualities, showing that high-quality new products can improve brand reputation and thus increase existing products' sales (backward spillovers).

Our paper is most closely related to that of Hendricks and Sorensen (2009), who made the first empirical contribution on information spillovers between products. They studied information spillovers between music albums in the pre-digital US music market, from 1993 to 2002. They show there is a substantial and persistent increase in sales of an artist's catalog albums due to discovery during the release of an artist's new album. In particular, backward spillovers created by the release of a second album increase first album sales from 40 to 55%.

In contrast to their paper, we study information spillovers in the digital age, between 2003 and 2010, in the French music market. We estimate that backward spillovers increase first album sales on average by 30%, but our main contribution is to evaluate the impact of digitization on the magnitude of information spillovers, between 2003 and 2007.

Our paper is related to the academic literature on the "Long Tail" hypothesis, a term introduced by Anderson (2004) to describe how digitization and online distribution may allow niche products to make up larger share of total sales than they would in the pre-digital age. Anderson (2006) argues that the exposure to a greater variety of products leads consumers to

streaming services LastFm (2006) and Deezer (2007). The development of online recommendation tools is also illustrated in Figure 1 in Appendix.

[10] In 2012, the number of Unique Monthly French Visitors was (in millions): 6 to SkyBlog, 10.4 to Blogger, 28.4 to Facebook, 3 to MySpace, 23.5 to YouTube, 11 to Dailymotion, 5 to Deezer, 3 to Twitter (ComScore/Médiamétrie/IDATE).

find the variety closest to their most preferred choice. But, empirical studies provide mixed evidence about the existence and the magnitude of the Long Tail (Brynjolfsson *et al.*, 2003; Anderson, 2006; Elberse and Oberholzer-Gee, 2007; Tucker and Zhang, 2007). Some argues that online recommendation tools (i.e., online communities, social networks, and personalization and recommendation technologies) should allow consumers to easily acquire product information with greater convenience and at lower costs, leading to increased demand for less popular products (Anderson, 2006; Brynjolfsson *et al.*, 2006).

Hervas-Drane (2012) use a theoretical model to show that recommender systems based on social filtering alongside traditional word-of-mouth recommendations have a positive impact on consumers interested in niche products, since such recommenders are more likely to draw attention to niche products. Oestreicher-Singer and Sundararajan (2012) show that recommendation systems that create hyperlinked content network, like on Amazon, cause product sales to be more evenly distributed. As regards exchange communities, Peitz and Waelbroeck, (2006) underline that consumers can discover new products with peer-to-peer technologies; this “consumer sampling” may replace costly marketing and promotion, and seems to prevail for lesser-known artists (Gopal *et al.*, 2006). Similarly, consumers’ awareness about artists ignored by mass media could increase as artists use online promotion tools (e.g., MySpace, Facebook, or YouTube) to obtain visibility (Bastard et al., 2012).

On the contrary, online recommendation tools could increase the informational inequality between hit and niche products. Fleder and Hosanagar (2009) use theoretical models and simulation to predict that recommendation systems that base on sales and ratings reinforce the popularity of already popular products. They show also that recommendation systems could increase individual-level diversity, but reduce aggregate diversity: Besides, consumers’ reviews appear even more skewed towards popular products online than offline, both for movies (Dellarocas et al., 2010) and music artists (Bastard et al., 2012). With an experimental design, Salganik et al. (2006) study the effect of the availability of popularity information on consumer decisions in an artificial online music market, and find that download counts reinforce consumer interest in the most popular products.

Within this literature, our paper contributes to the ongoing debate about the extent to which the development of online recommendation tools could increase information about “long tail products”. But in this paper, we try to assess the overall impact of online recommendation tools using information spillovers.

More broadly, we contribute to the growing literature about the impact of information provision on market outcomes in cultural industries. Previous studies have found that consumer choices can rely on advertising and promotion (Prag and Casavant, 1994), on prices and awards received (Litman, 1983), on expert reviews (Reinstein and Snyder, 2005), on best-seller lists (Sorensen, 2007), or on word-of-mouth [11] (Arndt, 1967). In the digital age, researchers have especially studied the impact of consumers' reviews and ratings (Senecal and Nantel, 2004; Chevalier and Mayzlin, 2006), online rankings (Salganik et al., 2006), peer-to-peer technologies (Peitz and Waelbroeck, 2006; Gopal *et al*., 2006), and recommendation systems (Fleder and Hosanagar, 2009; Hervas-Drane, 2012; Oestreicher-Singer and Sundararajan, 2012).

3. Data and descriptive statistics

3.1 Data

We construct two samples of music artists in order to compare the magnitude of backward spillovers between a first sample of artists who debuted in 2003, and a second sample of artists who debuted in 2007, when online recommendation tools had become more widely available.

We use a data set of weekly physical album sales in France between 2003 and 2010.[12] Music sales are tracked at the point of sale by monitoring cash registers at over 3,500 retail outlets. The panel of retailers is representative of album sales in France, and is composed of various distribution channels, including supermarkets specialized in cultural products, food stores, record shops, online retailers, and other specialized stores. We observe weekly sales for each album from the time of its release through the end of 2010, and for each album, the artist name, the genre, the number of units sold and the sales.

As we wish to focus on music albums, we exclude film soundtracks, recordings of comedy shows, children's stories and audio books. Also, we restrict our attention to full-length studio releases and exclude singles, EPs, maxis, recordings of live performances, holiday albums, anthologies and compilations.

[11] Word-of-mouth is defined by Arndt (1967) as an oral form of interpersonal non-commercial communication among acquaintances.

[12] Data have been retrieved from the GfK Marketing Institute, which is the principal source of sales data for the French music industry and the basis for national charts and rankings of artists popularity.

We wish to identify artists who debuted in 2003 or 2007. Unfortunately, no charts exist in France that list upcoming new artists.[13] Therefore, we select from the whole weekly sales album database a set of artists who match our conditions, and then examine their discographies. In presenting the construction of the samples, we focus on artists who debuted in 2003 (the same method was applied for artists who debuted in 2007).

We keep all the albums for which the first weekly sale date happened in 2003, and for which the artists did not appear in the database before 2003 to ensure they made their first record in 2003. Then, we exclude artists who did not release another album. Lastly, since a lot of artists have sold a few units of their debut album, we restrict our attention to the artists who sold at least 1,000 units of their debut album to avoid a lot of weekly sales at zero.

We obtain 145 artists for 2003 and 127 artists for 2007. For each of these artists, we used various online databases to obtain the artists' discographies and the exact release dates of their first and second albums.[14] After dropping a small number of artists who did not release a second full-length album, and others who had already released an album before 2003 or 2007, we obtain a sample of 145 new artists for 2003 and 127 new artists for 2007, who sold more than 1,000 units of their first album and who released at least one other album.

3.2 Summary statistics

The descriptive statistics of our sample underline some relevant characteristics which we will take into account in our econometric model, such as the diversity of artists in the sample and the path of their first album sales.

Artists in the sample belong to six genres of music (see Table 1 in Appendix): Pop Rock (40%), Hip Hop-Soul-R&B (19%), French variety (16%), Electronic (9%), Jazz -Blues (7%), Classical (5%) and World Music (4%). Around 50% of the artists are French,[15] 25% come from America and 25% from the European Union (see Table 2 in Appendix).

[13] Hendricks and Sorensen (2009) used the "Heatseekers Billboard charts", which is the weeks' top-selling albums by new or developing acts, defined as those who have never appeared on the top 100 of the Billboard 200 or the top 10 of R&B/Hip-Hop Albums, Country Albums, Latin Albums, Christian Albums, or Gospel Albums.

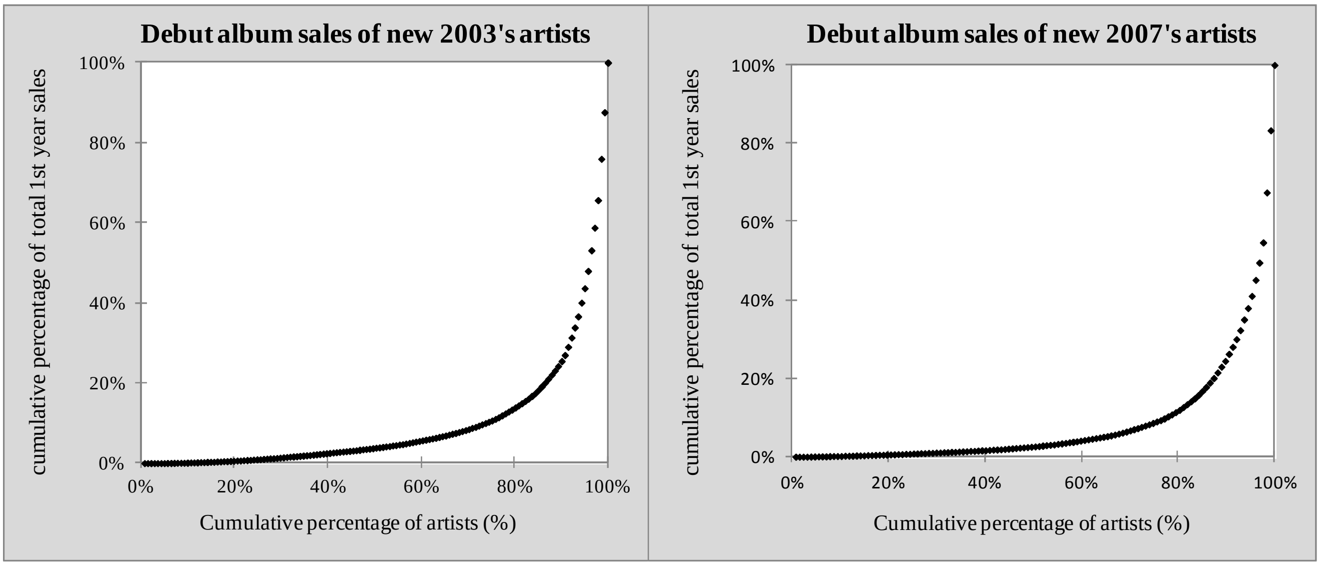
[14] We used five different online sources: an online database of information about audio recordings (*Discogs)*, an online music guide service (*AllMusic)*, the music website (*LastFm)*, *Wikipedia*, and the website of a French retail chain specialized in cultural products (*Fnac)*.

[15] The French music industry benefits from trade and government support, especially for the national live scene and non mainstream music. Also, radio stations must abide by the Domestic Quota which aims at encouraging investments in national production and strengthening the local market.

Some artists have contracted with major or independent labels, and they cover a broad range of commercial success: from the most successful like American metal group *Evanescence* or French singer *Thierry Amiel* in 2003, and international pop singer *MIKA* or French R&B group *Tragedie* in 2007, to relatively unknown and obscure artists like French singer *Anis* or Irish rock band *The Thrills* in 2003, and Australian electronic group *Midnight Juggernauts* or English soul band *Belleruche* in 2007.

We observe a strong heterogeneity in sales across albums, and a concentration on a few successful artists. Figure 1 shows the distribution of total first year sales across artists in each sub-sample. We observe that around 20% of the artists in both samples account for almost 85% of the first year sales. Median sales are about 4,300 units and average sales of 34,000 units, with a maximum of roughly 900,000 units. About a quarter of the artists sell between 5,000 and 17, 000 units, and only 10% sell more than 50,000 units during the first year following the release of their debut album.[16]

Figure 1. Distribution of First Year Debut Album Sales



The analysis of the sales shows that a large proportion of sales for the debut album occur during the first year following the release: on average, about two-thirds of sales are realized over the first year following the release. Most albums' sales paths exhibit an early

[16] According to SNEP (Syndicat National de l'Edition Phonographique), a commercial success in France is equivalent to a "gold record" certification (100,000 units sold, in 2003) or a "silver record" certification (50,000 units sold, in 2006). Sales between 10,000 and 50,000 characterize a medium commercial success, while sales between 1,000 and 5,000 copies may reveal a critical success. The crisis of the music industry led to a subsequent review of the level for a gold record certification (75,000 units, from May 1st, 2005) and silver record certification (35,000 units, from May 1st, 2005).

peak followed by a steady decline. Table 1 below displays the distribution of the week of first albums' sales peaks across the artists. During the first year following the release, sales peak on average during the 12th week. More precisely, 75% of the artists in both samples exhibit peak sales within the first 16 weeks, and only 10% after more than 35 weeks.

Table 1. Week of Peak Sales

Peak sales week								
	N	MEAN	ST. DEV.	0.10	0.25	0.50	0.75	0.90
2003 Sample	*145*	11.27	13.72	1	2	4	14	35
2007 Sample	*127*	11.89	13.39	2	3	6	16	33

We also need to consider the seasonality of albums sales and release dates. As expected, we observe that album sales are highly seasonal, like release dates: sales are strongest in spring and fall, and there is a huge increase in December in which sales are two to three times larger than average sales in other months of the year. Table 2 below shows the distribution of releases across months. Spring and fall appear to be the most popular periods to release a new album, whereas labels seem to avoid releasing new albums during summer and in December or January.

Table 2. Seasonal Variation of Release Dates

PERCENTAGE OF RELEASES OCCURING					
MONTH	2003-Sample		2007-Sample		Overall *(N=269)*
	Album 1	Album 2	Album 1	Album 2	
1 - January	6.4	4.1	8.7	4.5	5.9
2 - February	9.2	6.9	7.8	7.1	7.8
3 - March	9.9	13.1	12.6	12.5	12.0
4 - April	4.9	8.3	5.5	12.5	7.8
5 - May	7.8	11.0	11.8	8.0	9.7
6 - June	10.6	7.6	3.9	6.2	7.1
7 - July	9.2	4.1	2.4	1.8	4.4
8 - August	3.5	6.2	6.3	8.9	6.2
9 - September	8.5	10.3	7.1	13.4	9.8
10 - October	12.1	10.3	24.4	12.5	14.8
11 - November	13.5	14.5	7.9	8.9	11.2
12 - December	4.3	3.4	1.6	3.6	3.2

Table 3 summarizes some statistics for the album releases in our two samples: (a) the distribution of album's release dates, and (b) the delay between the releases.

Table 3. Album Release Statistics

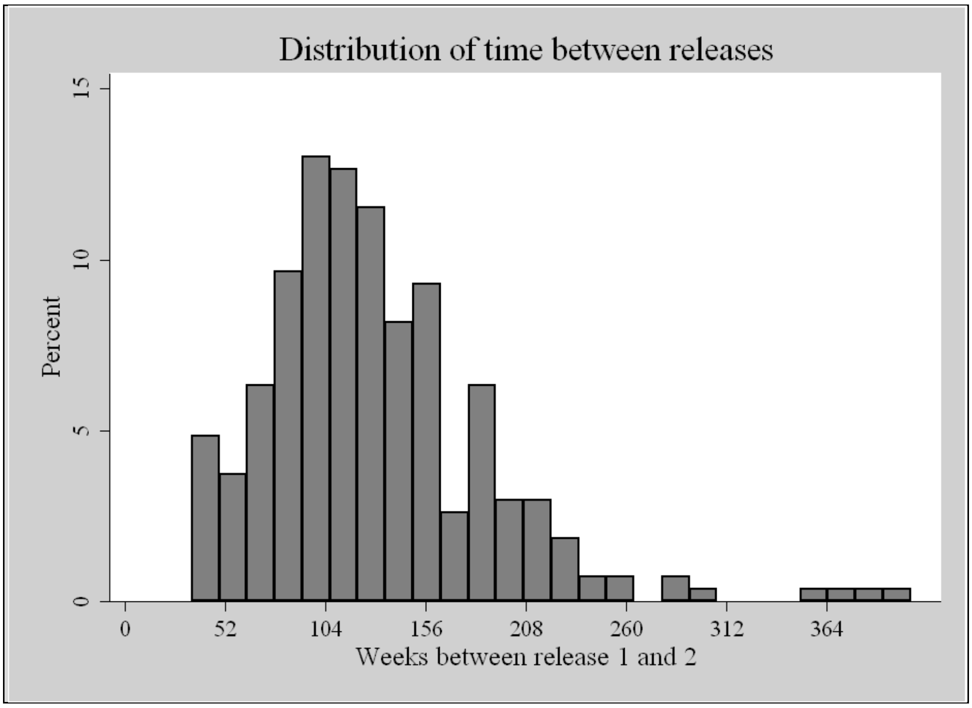
ALBUM RELEASES STATISTICS										
					PERCENTILE					
		N	MEAN	ST DEV.	0.10	0.25	0.50	0.75	0.90
a. Date Of Release									
Album 1	*2003 Sample*	*145*	25-Jun-03		17-Feb-03	17-Mar-03	23-Jun-03	20-Sep-03	27-Oct-03
	2007 Sample	*127*	30-Jun-07		5-Feb-07	19-Mar-07	9-Jul-07	1-Oct-07	12-Nov-07
Album 2	*2003 Sample*	*145*	13-Feb-06		14-Sep-04	6-Mar-05	1-Nov-05	14-Oct-06	18-Oct-07
	2007 Sample	*127*	11-Dec-09		14-Nov-08	20-Apr-09	2-Nov-09	7-Jul-10	14-Mar-11
b. Weeks Between Releases									
Album 1→2	*2003 Sample*	*145*	134.6	69.0	66	90	113	163	221
	2007 Sample	*127*	127.1	42.6	75	102	124	152	188
	Overall	*269*	131.0	58.8	66	94	120	158	205

The first part (a) of Table 3 shows the distribution of albums' release dates. The median debut date for artists in the 2003 Sample is June 23, and July 9 for artists in the 2007 Sample. For artists with a debut album in 2003, the median second album release date is November 1, 2005, with a mean at February 13, 2006. Some artists released their second albums as early as September, 2004 and others as late as October, 2007. For new artists starting in 2007, the median second album release date is November 2, 2009, with a mean at December 11, 2009. Some released their second albums as early as November, 2008 and others as late as March, 2011.

The second part (b) of Table 3 displays the delay between the releases of first and second albums. Overall, the median elapsed time before the release of the second album is more than two years (120 weeks), and the low end of the distribution is more than one year (66 weeks).

Figure 2 below shows precisely the histogram distribution of the lags between first and second album. Around 80% of the artists exhibit an elapsed time under 3.5 years (186 weeks) between the first and the second album release.

Figure 2. Distribution of Release Time (Histogram)



To illustrating backward spillovers, we show them graphically for two artists from the 2003 Sample, in Figure 3 and Figure 4 below. These two graphs plot the logarithm of weekly total sales over time for the artists' first album from the time of the artists' debut release, and the dashed vertical lines indicate the date of the release of their second albums.

In Figure 3, which follows the first album sales of the soul singer Amy Winehouse, we observe a strong impact of her hit second album "Black to Black", released in October, 2006 (week 157), on the sales of her debut album "Stronger than me", which failed to gain success in France at first. First album sales increase dramatically after the new release and peak almost one year later, remaining at a higher level than before releasing her second album.

As Figure 4 shows, the first album sales of the French singer Thierry Amiel, runner-up on TV show Pop Idol in France, exhibit a very different path. His debut album "Paradoxes" was an early success and awarded a gold record in France a few weeks after being released. Sales of the first album reached their peak in the early weeks following the release, and started decreasing from this point until the release of his second album "Thierry Amiel". In the weeks surrounding the second release, sales of the debut album experienced a surge for one year.

Figure 3. Debut Album Sales for Amy Winehouse

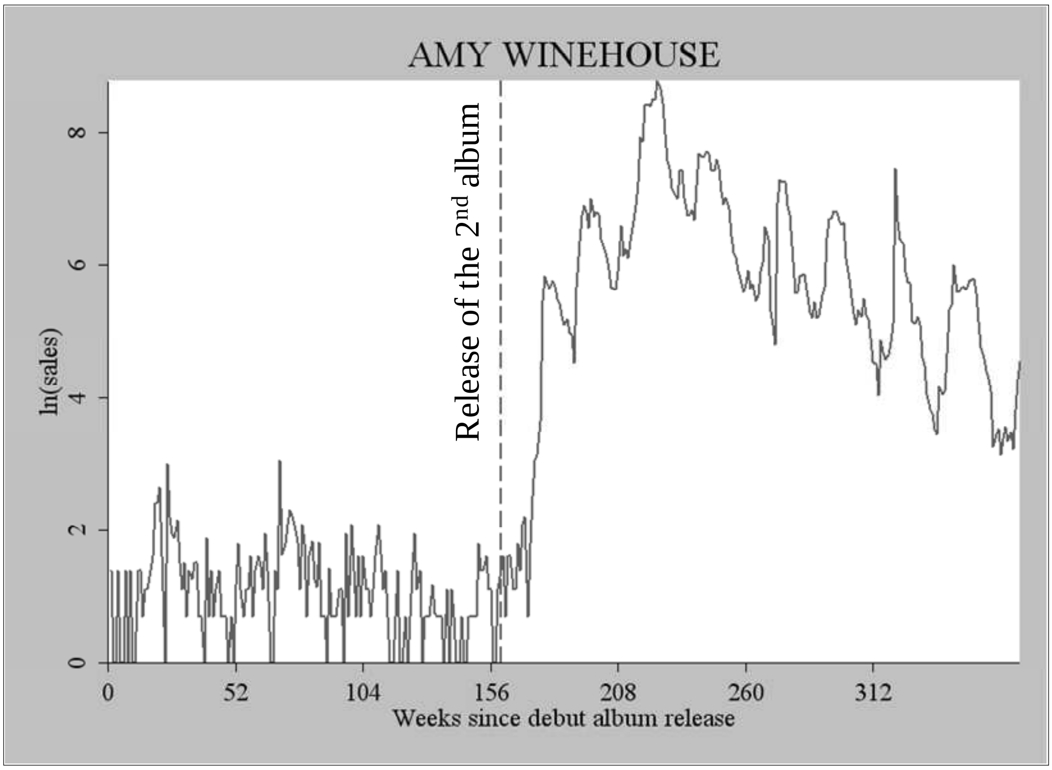
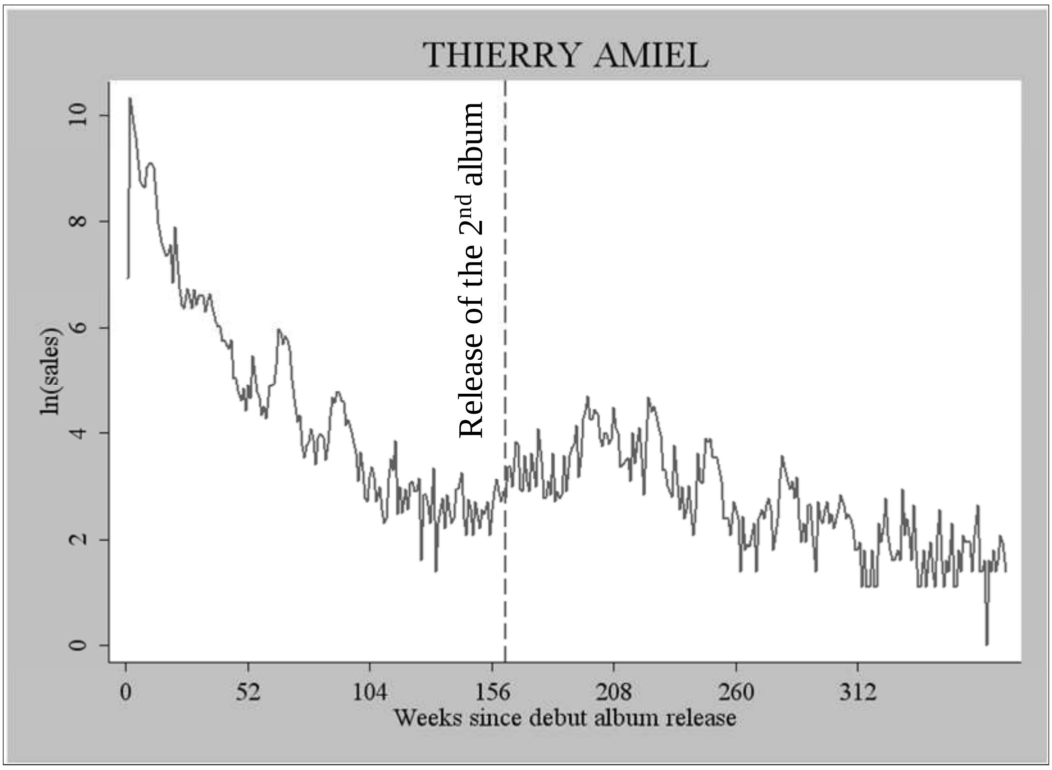


Figure 4. Debut Album Sales for Thierry Amiel



4. Empirical Strategy

To estimate backward spillovers, we follow the methodology proposed by Hendricks and Sorensen (2009). We observe the flow of sales for prior albums at the time when a new album is released, and both cross-sectional and time-series variations can be used to measure the sales response. The release of a second album by an artist represents the "treatment" on the first album's sales, and the treatment is an irreversible act.

We define S as a binary treatment indicator, where $S = 1$ with treatment and $S = 0$ without treatment. Our aim is to estimate the effect of the treatment, that is, the second album's release on the sales of the debut album, during *s* periods of treatment. We estimate the average treatment effect on the population treated for each period of the treatment window (*S*), which is the difference: $y_{it}^{s} - y_{it}^{0}$, where y_{it}^{s} is the sales of the first album of artist *i* in weekly period *t* in week *s* of the treatment window *S*, and y_{it}^{0} is the sales of the first album of artist *i* in weekly period *t* without treatment.

The treatment effect is the difference between two potential outcomes: the potential outcome with treatment for the treated album sales (that we observe) and the potential outcome for the album sales without treatment. The problem is that we do not observe this last outcome, the sales of an album in the absence of treatment. To estimate the counterfactual sales, we use albums that have not yet been treated as a control group, by exploiting the exogenous variation between release dates of albums.

Two key assumptions in the model also need to be underlined. First, we assume that prices are constant over time and across albums. This assumption is valid when we check the price of the first album over the period: if the first album can be discounted, it is not systemically related to the period surrounding the second release of the artist. Second, we assume that the preferences are additive across albums by the same artist, so that there is no complementarities in consumption between the first and the second album.

Following this approach, we run the following regression for our sample:

$$y_{it} = \alpha_0 + \alpha_i + \lambda_t + \sum_{m=2}^{12} \gamma_m D_{it}^{m} + \sum_{s=-5}^{26} S_{it}^{s} \,.\, (\beta_s + \delta_s W_i) + \epsilon_{it}$$

where y_{it} is the log of album sales of artist *i* at time *t*, and *t* is the number of weeks since the first album's release. The dependent variable is log-transformed to handle the positive skewness in the sales distribution.

α_i is an artist fixed effect which does not vary over time. The fixed-effects model controls for all time-invariant differences between individuals, so the estimated coefficients of the fixed-effects models cannot be biased because of omitted time-invariant characteristics. We assume that it is the time-invariant effect that impacts the treatment indicator, and not the idiosyncratic shock of the time-varying error term. Indeed, we need the treatment to be random across artists to ensure that the estimation of the average treatment is still valid. Like Hendricks and Sorensen (2009), we assume that the main determinant of the length of time between album releases is essentially an artist's creativity and personal effort, which are time-invariant and controlled by the artist fixed effect.

λ_t's are time dummies that control for the decay of the path of sales, and $D^{m\prime}$s are month dummies that control for seasonality.

S_{it}^{s} is an indicator equal to one if the release of artist *i*' s second album is *s* weeks away from period *t*, therefore β_s measures the second album's release impact on the artist's first album sales in week *s* of the treatment window. We allowed for a 32-week treatment window (i.e., 7 months), beginning 5 weeks (i.e., 1 month) before the week of the new release (in $t = 0$) and ending 26 weeks (i.e., 6 months) after the second release. The pre-release period allows us to estimate the effect of promotional activities done before the new release.

W_i is a dummy variable, equal to 0 if the artist's entry into the music industry was in 2003, and equal to 1 if the artist's debut was in 2007, when online recommendation tools had become more widely available. The interaction term between dummies *W* and *S* test the impact of the second album's release on first album sales depends on *W*, so δ_s measures a change in the coefficient of the treatment effect *S* over *W*. In other words, δ_s measure the variation in magnitude of the backward spillover between 2003 and 2007. We assume that a negative variation (i.e., a decrease of the backward spillovers between 2003 and 2007) imply that consumer information have increased as online recommendation tools become more widely available.

We include in the sample for each *t* artists who have released their second album and artists who have not yet released one, and exclude artists whose catalogs have been already treated by releasing a second album. We start to include debut album sales at *t* equal to 35 weeks (i.e., 8 months) to ensure we do not model an early peak in album sales and that the λ_t's better control for the time decay dynamic. We stop including albums at *s* equal to 17 in order to eliminate post-estimation treatment in our regression.

After performing a Breush-Pagan test and a modified Wald test, we corrected standard errors to take into account heterosckedasticity across individuals, because some artists' sales are more volatile than others. We also corrected for serial correlation within individuals, after detecting auto-correlation of the stochastic errors of the first order auto-regressive form. So the $\hat{\rho}$'s are the estimated AR(1) coefficients, reflecting the degree of serial correlation in demand shocks for a given album.

5. Results

Table 3 in the Appendix, reports the estimation results. The rows list the estimated coefficients β_s and δ_s for the 32 weeks of the treatment window. Time and month dummies were included in the regressions but estimated coefficients are not reported to facilitate reading.[17]

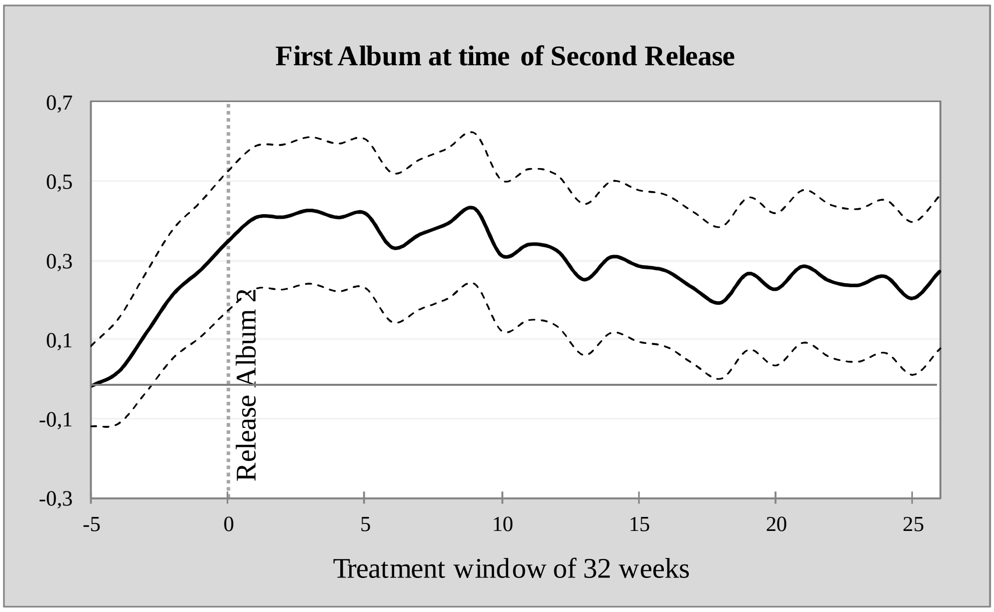
Figure 5 below summarize the results graphically, showing the estimated effect along with 95% confidence interval bands.

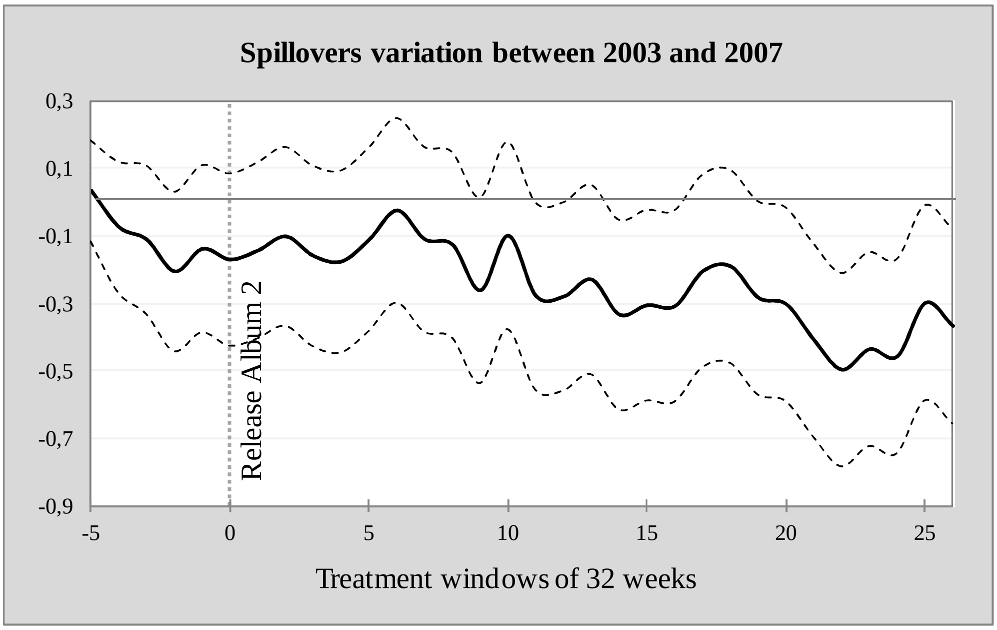
The first figure, entitled "First album sales at time of second release", shows the estimated effects of the second album release on the sales of the debut albums. Since the dependent variable is the logarithm of sales, the coefficients can be interpreted as percentage changes in sales of the first album resulting from the second release. In general, small (but statistically significant) increases start showing up 2 weeks prior to the new album's release (which occurs at $t=0$), growing in magnitude until 5 weeks after the release (which occurs at $t=5$). Overall, backward spillovers are between 12% and 43%, on average about 31%, and their effects for each of the weeks following the release of the second album are always positive and statistically significant.

The second figure, entitled "Spillover variation between 2003 and 2007", shows the estimated coefficients of the interaction terms between the 32 weeks of the treatment window and the indicator variable W of digital development. We find a significant and negative statistical effect of the spillover variation between 2003 and 2007 for the 15 weeks between the 9th and the 26th week of the treatment window, and all the coefficients are negative in each period of the treatment window. On average, the decrease of backward spillovers is of about 34%.

[17] They appeared all statistically significant, and time dummies t reveal a steady and monotonically decline over time.

Figure 5. Backward Spillovers





We interpret our results by considering the different process by which consumers obtain the knowledge of an artist, since both mass media visibility and word-of-mouth contribute to consumer information. Word-of-mouth is one of the three main sources of influence to purchase music, together with radio and television (Peitz and Waelbroeck, 2005).

The academic literature distinguishes two stages of information dissemination in the diffusion of new products in a market. The literature on the diffusion of new products [18]

[18] based on the framework developed by Bass (1969), A new product growth model for consumer durables. *Management Science* 15 (5): p215–227.

considers that new adopters join the market as a result of two types of influence: external influences, such as advertising and other communications initiated by the firm, and internal influences that result from interactions among adopters and potential adopters, in terms of word-of-mouth and personal communications.

In the period surrounding a new album's release, which we call the "traditional promotion window", mass media visibility enhances consumers' awareness about the artist. This traditional form of promotion includes advertising and marketing expenditure, mass media cover, radio airplay and television broadcasting as well. In a second period, which we refer to as the "word-of-mouth window", consumers' awareness about the artist may increase as word-of-mouth spreads about the new release; therefore, backward spillovers will occur later after the second release.

We are not aware of any research that would have estimated the timing of word-of-mouth and mass media promotion. However, promotion efforts and advertising should be mainly focuses around the release date of an album, as radio plays and video clips on television. Goldenberg et al. (2001) show that, beyond a relatively early stage (i.e. 16% of the market becomes informed), the effect of external marketing efforts or advertising quickly diminishes and word-of-mouth becomes the main factor driving the diffusion of new products. We find evidence that word-of-mouth takes time to spread information and to become influential in empirical paper. Leskovec et al. (2007) show that the probability of purchasing music dramatically increases when an individual has received recommendations from at least five other individuals in his networks, and those recommendations are more efficient for cultural niche products.

Consequently, we can reasonably infer that the "traditional promotion window" occurs about one month prior and after the release of a new album, while the "word-of-mouth window" occurs as of one month after the new release as word-of-mouth takes time to spread information. Building on this idea, we perform test for joint significance in the two different period windows. Our results are summarized in Table 7 below.

Table 7. Test for Joint Significance

Windows	Weeks	F-Test	Effect of digital technologies on consumer information
Full window (32 weeks)	From 5th week pre-release to 26th week post-release	$\chi2_{(32)}$ = 40.59 p-value = 0.14	**None**
Traditionnal promotion (9 weeks)	From 5th week pre-release to 4th week post-release	$\chi2_{(9)}$ = 6.33 p-value = 0.71	**None**
WOM window (23 weeks)	From 5th week post-release	$\chi^2_{(23)}$ = 35.77 p-value = 0.04	**Positive** (decrease of backward spillovers)

We find a statistically significant effect at the 5% level ($\chi^2_{(23)}$= 35.77, *p-value* = 0.04) when starting at fifth week following the second release, in the "word-of-mouth window". In contrast, we find no statistically significant effect at the 5% level ($\chi^2_{(23)}$= 6.33, *p-value* = 0.71) in the "traditional promotion window", as in the full window ($\chi^2_{(32)}$ = 40.59, *p-value* = 0.14).

According to our results, backward spillovers are the same in 2003 as they are in 2007 with respect to the "traditional promotion window". The impact of the second album's release on the artist's first album sales — because of consumers' discovery due to the promotion and mass media visibility of the second album — does not seem to have change over time with the development of online recommendation tools.

In contrast, backward spillovers are lower in 2007, as of one month after the second album's release, i.e., backward spillovers have decreased in the "WOM window". Our results suggest that word-of-mouth about the second album has a lower impact on the first album's sales for artists who debuted in 2007 than for those who debuted in 2003. We can assume that word-of-mouth is more widely diffused through the Internet and may improve consumer information by allowing more consumers to learn about the artist right after her first release.

Finally, the decrease of backward spillovers in the second stage of information dissemination suggests that digital technologies improve consumer information through the effects of online word-of-mouth. Indeed, online communities have been developed on the Internet, far beyond personal relationships, in which individuals share the experience they have of products and services among people of various backgrounds (Van Alstyne and

Brynjoflsson, 2005). Discussion forums, blogs, social networks, peer-to-peer technologies, or collaborative filtering systems, give an opportunity to extend word-of-mouth, even giving it "*a new significance due to the unique property of the Internet*" (Dellarocas, 2003 p. 1407).

6. Concluding Remarks

In this paper, we study how digital technologies have been impacting product discovery, in the demand for recorded music. In many markets, like the music market, consumers are exposed to an extensive supply of goods, and they are traditionally poorly informed about most products. We assess empirically whether consumers are better informed about products that fit their tastes as online recommendation tools – including online communities, social networks, or personalization and recommendation technologies [19] – become more widely available.

We compare the magnitude of "backward spillovers", originating from the empirical model of Hendricks and Sorensen (2009), to estimate consumers' lack of information about music albums at two different stages of the digital development: in the early stage, in 2003, and when online recommendation tools and platforms had become more widely available, in 2007.

We find that backward spillovers have significantly decreased between 2003 and 2007, in the second stage of information dissemination. Our results imply that online recommendation tools have been increasing consumer information, via *electronic word-of-mouth*. Online communities, social networks, and peer-to-peer technologies give an opportunity to extend and relay the word-of-mouth on an unprecedented scale. With the rise of social networks and online communities, the Internet could increase the number of "weak ties" [20] across individuals (Donath and Boyd, 2004), which in turn could rise the speed of information dissemination (Goldenberg et al., 2001).

A first remark is that our estimation of the backward spillover decrease may reflect the aggregation of two opposite effects. Indeed, we assumed that artists who debuted in 2007

[19] "Personalization and recommendation technologies take various forms, from simple top lists of recommended items other users have selected to sophisticated collaborative filters that infer preferences with little or no active effort by the consumer" (Brynjofsson et al., 2010).

[20] Weak ties are acquaintances or loose relationships; they act as bridge links and are more effective at disseminating information because the information they transmit to one another is more likely to be new. On the other hand, strong ties, often defined as close friends and family, usually possess similar information and this limits the sharing of new information (Granovetter, 1983).

benefit from better information dissemination than artists who debuted in 2003. Indeed, more consumers know about the artist right after his first release, leading to a decrease of backward spillover. However, a larger portion of the remaining uninformed consumers may learn about the artists at the time of the second release, thus leading to an increase in backward spillover. Overall, our estimation of backward spillover variation aggregates these two opposite effects and suggests that the first effect dominates.

Several limitations arise from our dataset. First, we use only weekly physical album sales, even from online retailers, but not online digital album sales. Second, sales from digital piracy are obviously “shadow” sales for which we also cannot account for. Although these missing sales could lead to under-estimating backward spillovers, some evidence mitigates this assumption. Market figures show that digital sales represent only from 7% to 15% of the music market between 2003 and 2010. Also, the effect of digital piracy on spillover variation could be limited: some studies show that pirates are also the ones who buy more cultural products (Bounie *et al.,* 2012) and that the “consumer sampling effect” of file-sharing technologies helps consumers to discover new products and improve consumer information (Peitz and Waelbroeck, 2006).

References

Adler, M. (1985). Stardom and Talent, *American Economic Review*, 75: 208-12.

Anderson C, (2006). *The Long Tail*, Hyperion press, New York.

Arndt, J. (1967). Role of product-related conversations in the diffusion of a new product, *Journal of Marketing Research*, 4: 291–5.

Bastard I., Bourreau M., Maillard S. et F. Moreau (2012). De la visibilité à l'audience : les musiciens sur Internet , Réseaux, 2012/5 (n° 175), p.19-p.42.

Bounie, D., Bourreau, M. and Waelbroeck, P. (2007). Pirates or Explorers? Analysis of Music Consumption in French Graduate Schools, *Brussels Economic Review*, 50(2): 167-192.

Brynjolfsson, E., Hu, Y.J. and Smith, M.D. (2003). Consumer Surplus in the Digital Economy: Estimating the Value of Increased Product Variety at Online Booksellers, *Management Science*, 49(11): 1580–1596.

Brynjolfsson, E., Hu, Y.J. and Smith, M.D. (2006). From niches to riches: The anatomy of the long tail, *Sloan Management Review*, 47(4): 67–71.

Brynjolfsson, E., Hu, Y.J. and Smith, M.D.(2010) Research Commentary - Long Tails vs. Superstars: The Effect of Information Technology on Product Variety and Sales Concentration Patterns. *Information Systerms Research*.

Cabral, L.M.B. (2000). Stretching Firm and Brand Reputation. *Rand Journal of Economics*, 31(4) 658-673.

Chevalier, J.A. and Mayzlin, D. (2006). The Effect of Word of Mouth on Sales: Online Book Reviews, *Journal of Marketing Research*, 43: 345–354.

Dellarocas, C. (2003). The Digitalization of Word of Mouth: Promise and Challenges of Online Feedback Mechanisms, *Management Science*, 49(10): 1407-1424.

Dellarocas, C., Gao, G. and Narayan, R. (2010). Are Consumers More Likely to Contribute Online Reviews for Hit Products or Niche Products? *Journal of Management Information Systems*, 27(2): 127-157.

Donath, J. and Boyd, D. (2004). Public Displays of Connection, *BT Technology Journal*, 22 (4): 71-82.

Elberse, A. and Oberholzer-Gee, F. (2007). Superstars and Underdogs: An Examination of The Long Tail Phenomenon in Video Sales, *Marketing Science*, 4: 49-72.

Fleder, D. and Hosanagar, K. (2009). Culture's Next Rise or Fall: The Impact of Recommender Systems on Sales Diversity, *Management Science*, 55(5): 697–712.

Goeree, M.S. (2008). Advertising in the U.S. Personal Computer Industry, *Econometrica*, 76(5): 1017-1074.

Goldenberg, J., Libai, B. and Muller, E. (2001). Talk of the network: A complex systems look at the underlying process of word-of-mouth, *Marketing Letters*, 12(3): 211–223.

Gopal, R., Bhattacharjee, S. and Sanders, G. L. (2006). Do Artists Benefit from Online Music Sharing? *Journal of Business*, 79: 1503-1534.

Granovetter, M. (1983). The Strength of Weak Ties: A Network Theory Revisited, *Sociological Theory*, 1: 201–233.

Hendricks, K. and Sorensen, A. (2009). Information and the Skewness of Music Sales, *Journal of Political Economy*, 117(2): 324-369.

Hervas-Drane, A. (2012). Search, Product Recommendations, and Sales Concentration. *NET Institute Working Paper No. 07-41*.

Katz, E. and Lazarsfeld, P. (1955), *Personal Influence*, New York: The Free Press.

Leskovec, J., Adamic, L. and Huberman, B. (2007). The Dynamics of Viral Marketing, *ACM Transactions on the Web (TWEB)*, 1(1).

Litman, B.R. (1983). Predicting the success of theatrical movies: An empirical study, *Journal of Popular Culture*, 17: 159-75.

Oestreicher-Singer, G. and Sundararajan, A. (2012). The Visible Hand of Social Networks in Electronic Markets, *Management Science*.

Peitz, M. and Waelbroeck, P. (2005). An Economist's Guide to Digital Music. *CESifo Economic*, 51(2-3): 359-428.

Peitz, M. and Waelbroeck, P. (2006). Why the Music Industry May Gain From Free downloading - The role of Sampling, *International Journal of Industrial Organization*, 24: 907-913.

Prag, J. and Casavant, J. (1994). An Empirical Study of Determinants of Revenues and Marketing Expenditures in the Motion Pictures Industry, *Journal of Cultural Economics*, 18(3): 217-235.

Reinstein, D. and C. Snyder. (2005). The Influence of Expert Reviews on Consumer Demand for Experience Goods: A Case Study of Movie Critics, *Journal of Industrial Economics*, 53: 27-51.

Rogers, E.M. (1976). New product adoption and diffusion, *Journal of Consumer Research*, 2: 290-301.

Rosen, S. (1981). The Economics of Superstars, *American Economic Review*, 71: 845-58.

Salganik, M., Dodds, P. and Watts, D. (2006). Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market, *Science*, 311: 854–56.

Sénécal, S. and Nantel, J. (2004). The Influence of Online Product Recommendations on Consumers' Online Choices, *Journal of Retailing*, 80(2): 159-169.

Sorensen, A. T. (2007). Bestseller Lists and Product Variety, *The Journal of Industrial Economics*, 55: 715–738.

Tucker, C. and Zhang, J. (2007). Long tail or steep tail: A field investigation into how popularity information affects the distribution of customer choices. *MIT Sloan School, Working Paper 4655-07*, Cambridge, MA.

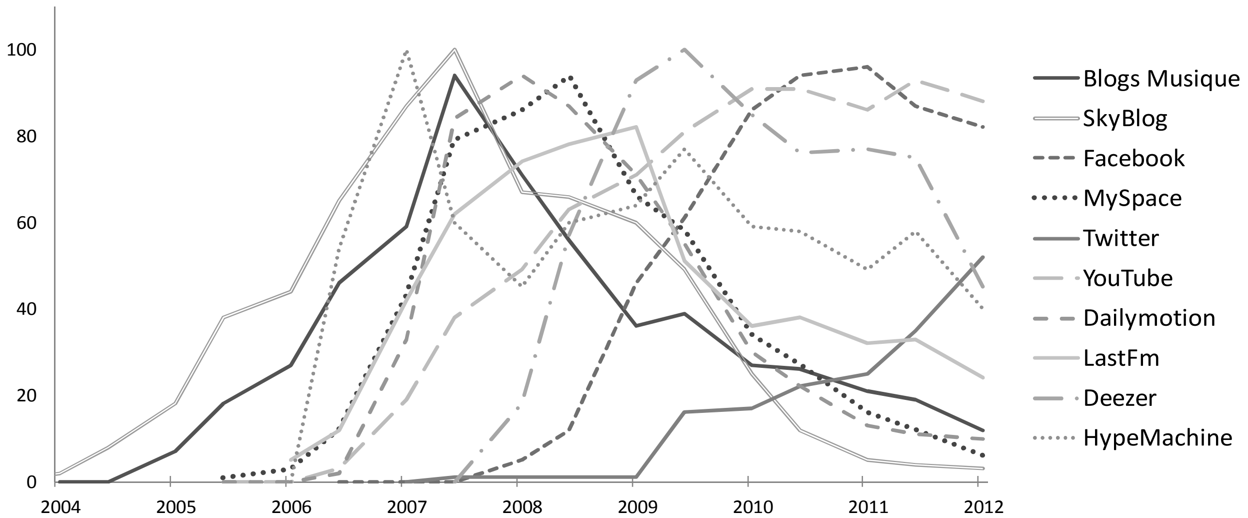
Van Alstyne, M. and Brynjolfsson, E. (2005). Global Village or Cyberbalkans: Modeling and Measuring the Integration of Electronic Communities, *Management Science*, 51(6): 851-868.

Wooldridge, J. (2002). *Econometric Analysis of Cross-Section and Panel Data*, Cambridge, MA: MIT Press.

Appendix

Figure 1. The development of online recommendation tools

Evolution of Google search terms in France between 2004 and 2010 (*source: Google Trends)*



Note:
Google Trends designates a certain threshold of traffic for search terms, so that those with low volume won't appear and are equal to zero . The graph reflects how many searches have been done for a particular term, relative to the total number of searches done on Google over time. Data are normalized and presented on a scale from 0-100.

Figure 2. Launch date of online platforms

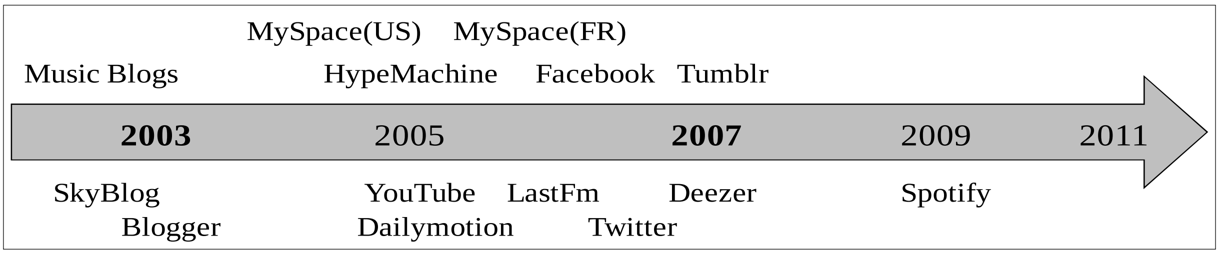


Table 1. Distribution of musical genres (in %)

Genre	2003-Sample *(N=145)*	2007-Sample *(N=127)*	Overall *(N=272)*
Pop Rock	32.4	47.2	39.8
Hip Hop RnB Soul	17.9	20.5	19.2
French variety	22.8	9.4	16.1
Electronic	9.7	8.7	9.2
Jazz Blues	8.3	6.3	7.3
Classical	4.8	4.7	4.8
World Music	4.1	3.1	3.6
Total	100	100	100

Table 2. Distribution of artists' origin

Origin	2003-Sample *(N=145)*	2007-Sample *(N=127)*	Overall *(N=272)*
France	50%	42%	46%
America	32%	21%	26%
European Union	14%	30%	22%
Others	5%	7%	6%
Total	100%	100%	100%

Table 3. Estimation results (treatment window of 32 weeks)

Week relative to release date of second album *(t=0)*	β_s	δ_s
t=-5	*-0.018* *(0.725)*	*0.302* *(0.667)*
t=-4	*0.019* *(0.778)*	*-0.075* *(0.448)*
t=-3	*0.116* *(0.130)*	*- 0.112* *(0.317)*
t=-2	*0.216**** *(0.009)*	*- 0.206* *(0.087)*
t=-1	*0.277**** *(0.001)*	*-0.140* *(0.269)*
t=0	*0.349**** *(0.000)*	*-0.171* *(0.189)*
t=1	*0.408**** *(0.000)*	*-0.144* *(0.278)*
t=2	*0.409**** *(0.000)*	*-0.103* *(0.447)*
t=3	*0.426**** *(0.000)*	*-0.161* *(0.239)*
t=4	*0.408**** *(0.000)*	*-0.177* *(0.198)*
t=5	*0.419**** *(0.000)*	*-0.112* *(0.419)*
t=6	*0.332**** *(0.001)*	*-0.026* *(0.852)*
t=7	*0.365**** *(0.000)*	*-0.111* *(0.426)*
t=8	*0.392**** *(0.000)*	*-0.127* *(0.364)*
t=9	*0.431**** *(0.000)*	*-0.262** *(0.062)*
t=10	*0.311**** *(0.001)*	*-0.100* *(0.478)*
t=11	*0.340**** *(0.000)*	*-0.279*** *(0.050)*
t=12	*0.325**** *(0.001)*	*-0.281*** *(0.049)*
t=13	*0.251**** *(0.010)*	*-0.230** *(0.100)*
t=14	*0.308**** *(0.002)*	*-0.334*** *(0.020)*
t=15	*0.285**** *(0.003)*	*-0.306*** *(0.034)*
t=16	*0.273**** *(0.005)*	*-0.309*** *(0.033)*
t=17	*0.229*** *(0.019)*	*-0.206* *(0.158)*
t=18	*0.192*** *(0.049)*	*-0.191* *(0.191)*
t=19	*0.266**** *(0.007)*	*-0.284*** *(0.050)*

t=20	*0.226*** *(0.021)*	*-0.303*** *(0.039)*
t=21	*0.284**** *(0.004)*	*-0.409**** *(0.005)*
t=22	*0.247*** *(0.012)*	*-0.497**** *(0.001)*
t=23	*0.236*** *(0.016)*	*-0.436**** *(0.003)*
t=24	*0.259**** *(0.008)*	*-0.457**** *(0.002)*
t=25	*0.203*** *(0.039)*	*-0.300*** *(0.043)*
t=26	*0.271**** *(0.006)*	*-0.367*** *(0.013)*
# of artists	268	
# of observations	31,835	
$\hat{\rho}$	0.827	

Model estimated with GLS, corrected for heterosckedasticity and serial correlation (AR1). Standard errors are in parentheses; *,**, and *** denote significance at the 10%, 5%, and 1% level, respectively.