Interdisciplinary Institute for Innovation

"Selling Less of More"? The Impact of Digitization on Record Companiess

Marc Bourreau
Michel Gensollen
François Moreau
Patrick Waelbroeck

Working Paper 12-TS-03

February, 2012

Department of Economics and Social Sciences
TELECOM ParisTech
46 rue Barrault, 75013 Paris, France
Email: wpecon@telecom-paristech.fr

"Selling Less of More"? The Impact of Digitization on Record Companies

Marc Bourreau*, Michel Gensollen*,
François Moreau** and Patrick Waelbroeck*

* Telecom ParisTech, Department of Economics and Social Sciences, Paris, France
** ICI, Université de Bretagne Occidentale, Brest, France

February 2012

Abstract

In this paper we use data from a survey of 151 French record companies to test the "long tail theory" at the level of the firm. More specifically, we test whether, following the "selling less of more" principle coined by Anderson (2006), record companies that have adapted to digitization (at various levels: artists' scouting, distribution and promotion) release more new albums without having higher overall sales. We construct a production function in which the output is produced from conventional inputs of labor and capital, as well as inputs that are more specific to the recorded music industry. We consider two types of output: a commercial output (albums sales) and a creative output (number of new albums released). We show that labels that have adapted to digitization are more efficient in respect of creative output, but that there is no effect of adaptation to digitization on the commercial output, which is consistent with the predictions of the long tail theory.

Keywords: Recorded music industry; Digitization; Long tail; Innovation.

JEL codes: Z11; O33; L2; D2.

1. Introduction

The impact of digital technologies on the recorded music industry is ambiguous. On the one hand, digitization has made it possible to share and exchange music files, and the resulting steep increase in music downloads on peer-to-peer networks is today seen by major record companies as the cause of the downturn in music sales. On the other hand, digital technologies have also reduced record companies' costs for the production of music (e.g., because artists can record music with a home studio), its distribution (on digital platforms) and its promotion (owing to the development of online word-of-mouth, for example on social networks such as MySpace or Facebook).

Whereas the effect of peer-to-peer downloads on music sales has been a subject of keen interest in the past few years (e.g, see Liebowitz, 2008; Oberholzer-Gee and Strumpf, 2007),[1] the potential effects of digitization on record companies has received less attention. In this paper, our aim is to determine how digital technologies have affected the activity of record labels.

According to Anderson (2006), the digitization of content industries might erode the high concentration of sales experienced in these industries. A few hundred items usually account for the bulk of sales, whereas the products belonging to the tail of the distribution of sales sell only a few units per month or even per year. The argument of the so-called "long tail theory" is as follows. Digitization first leads to a drop in production costs, which lengthens the tail and thus brings more products onto the market. Secondly, distribution costs decrease as well, which contributes to flattening the tail by making it easier for niche products to enter the market. The third effect of digitization is the development of online word-of-mouth, which ensures a better match between supply and demand, and thus drives business from hits to niches. According to Anderson, these trends can be expected to progressively increase the sales of obscure artists, to the detriment of stars. As variety increases, it becomes easier for consumers to discover niche artists who really match their preferences.

[1] See also Belleflamme and Peitz (2010) for a recent survey on the economics of digital piracy.

Anderson (2006) notes that the tail, made up of hundreds of thousands of items, can represent as many sales as the few hits belonging to the head of the distribution. As he puts it, selling less (units) of more (products) might then be the future of business.

Some recent studies have tested the existence of a long tail phenomenon in online markets for various products: clothes (Brynjolfsson et al., 2011), videos (Elberse and Oberholzer-Gee, 2008; Kumar et al., 2011), books (Brynjolfsson et al., 2010; Peltier and Moreau, 2012), and music (Benghozi and Benhamou, 2010). These papers analyze whether the share of the sales of niche products compared to hit products is larger online than offline. However, to the best of our knowledge, no research has been conducted to assess the impact of digitization on the supply and the sales of record companies. This paper aims at filling this gap.

According to the long tail theory, record labels that have adapted to digitization should release more albums, targeting niche sub-markets that are ignored by hit products. However, as Anderson (2006) has pointed out, each of these new releases probably sells less than the average album. Therefore, labels that take the step of digitization and decide to expand their catalog might have lower aggregate sales than record companies that have kept the traditional business model of the industry.

To test these two claims, we use firm-level data from the French recorded music industry. We first test if French music labels that have adapted to digitization release more new albums than do other labels. We then test whether their album sales are lower than those of labels that have not adapted to digitization. Our estimation results suggest that, while digitization has a positive and significant impact on album releases, it has no effect on total record sales, at the firm level. In other words, we find evidence that digitized labels sell less (or as much as before) of more.

The rest of the paper is organized as follows. In Section 2 we briefly present the theoretical framework and our hypotheses. Section 3 introduces the specification of the econometric model and the estimation methods. Section 4 describes the data. In Section 5 we present the empirical results. Finally, Section 6 concludes.

2. Theoretical framework

We assume that the output of record companies is produced from conventional inputs of labor and capital, and can be described by a production function. Even in creative industries, as Gapinski (1984: 465) argues, "the relationship between inputs and output obeys a well-behaved, standard production format."

Throsby (2006) considers different outputs for different aspects of the arts' activities. In particular, he makes a distinction between "commercial" artistic output, which can be measured by audiences or sales, and "creative" artistic output, which can be measured by the number of creations (sculptures, paintings, compositions, etc.). We adopt Throsby's distinction and define two types of output for a record company: a *commercial output* (annual sales) and a *creative output* (annual number of albums released).

We thus assume that a record company *i* produces the two types of output, according to the following production functions:

$$y_i^{co} = f^{co}\left(A^{co}, K_i, L_i\right), \qquad (1a)$$

$$y_i^{cr} = f^{cr}\left(A^{cr}, K_i, L_i\right), \qquad (1b)$$

where y_i^{co} and y_i^{cr} denote the commercial and creative output, respectively, A^{co} and A^{cr} the technology for the commercial and creative output, respectively, K_i the capital input, and L_i the labor input.

Our general hypothesis is that labels that have adapted to digitization release more new albums than do other labels, but do not benefit from higher overall sales.

There are various reasons why digitization might lead to higher levels of output for record labels, for given levels of labor and capital.[2] First, the Internet facilitates the selection of new talent, as many artists are online today on specialized websites like MySpace Music, where

[2] For a general discussion on the effects of digitization, see for example Peitz and Waelbroeck (2005), Regner (2003), Easley et al. (2003), Premkuvar (2003), and Belleflamme and Peitz (2010).

they provide excerpts of their music. Second, the Internet lowers distribution costs, as the labels can now deliver their music on digital platforms, without going through a distributor. Third, decentralized promotion on the Internet might be both less costly and more efficient than traditional promotion, at least for some artists.

However, according to the long tail theory, digitization has a larger impact on creation (variety) than on the sales of individual products. Each new album is supposed to sell less than it used to, as total sales are now split over a larger number of works that match consumer preferences better. We therefore posit that:

Hypothesis 1: *Digitization does not positively impact the overall sales of record companies.*

Hypothesis 2: *Record companies that have adapted to digital technology release more new albums.*

3. Econometric models and estimation methods

Following various studies devoted to efficiency in creative industries,[3] we assume a Cobb-Douglas production technology for the commercial output:

$$y^{co} = A^{co} K^{\alpha} L^{\beta} e^{\varepsilon}, \qquad (2)$$

where y^{co} denotes the commercial output, A^{co} the technology for the commercial output, K the capital, L the labor, and ε is the error term. Parameters α and β denote the elasticities of commercial output with respect to capital and labor, respectively. By taking logarithms of equation (2), we obtain:

[3] Noteworthy references include Throsby (1977), who estimated a production function for the Australian performing arts, Gapinski (1980, 1984) who used US and British data to study the efficiency of performing arts firms (theatre, opera, symphony, ballet), Taalas (1997) who studied the production structure in Finnish theatres, and Bishop and Brand (2003) who built a production function for British museums in order to study their efficiency.

$$\ln(y^{co}) = \alpha \ln(K) + \beta \ln(L) + \ln(A^{co}) + \varepsilon. \quad (3)$$

We assume that ε is independent of $\ln(A^{co})$, $\ln(K)$ and $\ln(L)$. To control for endogeneity, we use control variables. We use OLS to estimate equation (3).

For the creative output, we choose a different specification. Two noteworthy features of the creative output are a relatively long tail and a strong mode at zero. Therefore, we use the Poisson regression model. We assume that

$$E[y^{cr}] = \exp(\lambda K + \gamma L + A^{cr} + \varepsilon'), \quad (4)$$

and estimate this model with a Poisson regression. However, as the Poisson model relies on the strong assumption that the conditional variance equals the conditional mean, we also estimate (4) with a binomial negative model.

Hypotheses 1 and 2 that we wish to test can now be specified as follows:

H0: $A^{co}_{adapted} = A^{co}_{non\ adapted}$.

Record companies that have adapted to the Internet do not sell more CDs than record companies that have not. The alternative is **H1:** $A^{co}_{adapted} > A^{co}_{non\ adapted}$. That is, with digitization, record labels sell more.

H0': $A^{cr}_{adapted} = A^{cr}_{non\ adapted}$.

Digitization does not affect the creativity of record companies. The alternative is **H1':** $A^{cr}_{adapted} > A^{cr}_{non\ adapted}$. That is, labels that have adopted the digital technology release more new albums than do others.

4. The data

The data were collected from a survey on French record companies that we conducted between July and October 2006. A questionnaire was mailed to an extensive list of 871 labels, both for-profit and not-for-profit.[4] A week later, all labels were contacted by phone to help them complete the questionnaire. Eventually, 187 labels agreed to answer the questionnaire, that is, 21.5% of the total number of labels on our list. However, we had to exclude 32 questionnaires, for two main reasons: i) some labels had just stopped their label activities; and ii) some questionnaires were not answered correctly. We also excluded four labels that had been founded at the time of the study, in 2006. We thus ended up with a final sample of 151 record companies. This sample contained record companies of different sizes: one major record company, a few large independent labels, and several very small labels.[5]

The questionnaire was composed of three main sections: a first section with general questions (name of the label, year founded, number of employees, etc.), a second section with questions pertaining to their label activities, and a third section in which we asked the labels for their opinion on the crisis faced by the music industry.

In what follows, we outline the specification of our variables. Table A1 in Appendix A gives the description of the variables, while Table B1 in Appendix B provides summary statistics.

4.1. Output variables

We specify the "commercial output" as the album sales in 2005 (in number of CDs sold) (*SALES*), and the "creative output" as the number of new albums released in 2005 (*ALBUMS*). As expected, both output variables are highly skewed. The median sales are 4,166 CDs sold,

[4] This extensive list of French record companies was compiled from the following professional directories: "L'Officiel de la Musique" (IRMA, 2006), and "Le Réseau", (IRMA, 2005) for popular genres of music; "Jazz de France", (IRMA, 2006) for jazz music; "Planète Musique", (IRMA, 2005) for traditional and world music; "Le guide du disque classique", (Editions Cité de la Musique, 2005) for classical music. This list corresponds – more or less–to the total population of French record companies (only some very small labels might be missing).

[5] Unfortunately, we cannot test whether our sample is representative of the population it was extracted from, as no variable (such as the number of employees) was available for the full population. However, the market share of independent labels in our sample (compared to the market share of the major) is similar to the market share of independent labels in the French market in 2006. The extrapolation of the number of albums sold in 2005 to the full population is moreover in line with the total number of albums sold that year in France.

whereas the average is 197,053. Only 19% of the labels in our sample sold more than 50,000 albums in 2005. The average number of album releases is 6.3, whereas the median is equal to 3. We find that 28.6% of the labels released zero or one album in 2005, while 16.3% released ten albums or more.

4.2. Input variables

(a) *Labor* is captured by two different variables: *EMPLOYEES* gives the number of employees in 2005 (excluding occasional employees and interns) and *PR%* represents the percentage of albums released by the record company which benefited from press relations that year.[6] This distinction enables us to take into account both a generic labor factor and a specific labor factor devoted to the promotion of new album releases. The firms in our sample are on average very small. Only 9% of the firms have 10 employees or more and 52% have zero or one employee.[7] This is consistent with the prevailing belief that the recorded music industry is an oligopoly with a competitive fringe. As for press relations, 33% of the labels in our sample had made a press relations effort for more than half of their albums.

(b) *Capital* is measured by two intangible capital variables: intellectual property and reputation. Intellectual property rights on recordings represent a very important asset for record companies, since such rights generate revenues not only from sales of CDs or digital files but also from airplay, music-video broadcasts on TV, etc. We measure the amount of music rights by the size of the catalogue in 2006 (*CATALOG*). The median size of a catalog is 14 albums, and 8% of the labels own 200 or more albums in their catalogue.

We proxy the reputation of a label with the number of demo tapes received per month in 2006 (*DEMOS*). A record company with a good reputation on the music scene can be expected to receive more demos than an unknown label. The number of demos appears to be highly skewed. Labels receive on average 30 music demos per month, but half of them receive 10 demos or less per month and 10% receive 100 demos or more per month.

[6] By including staff and adjuvants (promotional personnel, etc.) in the labor inputs, we follow Gapinski (1980, 1984).

[7] In our dataset, when a label has no employees, this actually means that only the entrepreneur is involved in the label's activity.

4.3. Control variables

We use the following control variables:

POPMUSIC: adaptation to digitization and output might differ across the music genres. In particular, record companies specialized in popular genres like pop-rock could be faster to adapt than labels in traditional genres like classical or jazz music, because their audience is more likely to be online. Therefore, we introduce a dummy variable, *POPMUSIC*, which states whether the label produces mainly popular music (pop-rock, soul, funk, R&B, rap, hip-hop, electro, techno, world, reggae, etc.). In our sample, 76% of the labels produce mainly popular music, whereas the rest of the labels produce jazz or classical music.

FORPROFIT: competition might lead for-profit labels to adapt faster than non-profit labels. Incentives for commercial success are also logically higher for for-profit organizations; hence, we expect larger sales with for-profit labels. 30% of the record companies that we interviewed were not-for-profit organizations.

MAJORDISTR. 10% of the labels in our sample were distributed by a major. These labels might benefit from a better distribution at retail stores, and therefore make higher sales. Commercial relations with a major might also favor adaptation to digitization, at least in some of its dimensions (e.g., for the incorporation of digital distribution in artists' contracts).

COMPIL. Compilations sell well. Hence, labels that have titles in compilations might obtain higher sales and might also release more albums. 57% of the labels belonging to our sample had titles in a compilation in 2005.

EUROPE. 77% of the labels are distributed not only in their national market (France), but also in other countries in Europe. We expect these labels to benefit from higher sales. A multimarket activity might also accelerate the adaptation to digitization, as a label might benefit from the experience of a more "advanced" market.

YOUNG. We specify as "young" the labels that were founded after 1999, that is, after the beginning of the Internet era. The age of the label has a priori an ambiguous effect on

adaptation to digitization. On the one hand, older labels might be more experienced than young ones, and thus be able to adapt faster. On the other hand, for young labels, the effects of digitization were probably taken into account in the definition of their initial business model. In our sample, 48% of the labels are "young".

4.3. Adaptation to digitization

Digitization has three main effects on the music industry. Firstly, it affects the distribution of music, with the development of digital music platforms like the iTunes Music Store. Secondly, it affects the promotion of new artists and new music. Digital technologies—and the Internet in particular—have fostered the development of decentralized promotion (i.e., word-of-mouth) for cultural products, to the detriment of traditional centralized promotion through the mass media.[8] Thirdly, talent scouting can benefit from digitization: digital technologies facilitate communication with artists, the quality of tape demos is improved with the development of home studios, and last but not least, websites like MySpace provide potential talents with the opportunity to advertise themselves more efficiently. Since digitization has different dimensions, labels may also adapt to digitization in different ways. We therefore consider different aspects of the adaptation to digitization.

One obvious first step for a record company is to set up a website. At the time of the study, 90% of the respondents had a website, most of which had been created around the year 2000 (the average year is 2001). In our analysis, the variable *AGEWEB* is equal to the number of years since the creation of the website (at the time of the study). It therefore captures whether the label is an "early" or a "late" entrant on the Web. However, the mere existence of a website does not mean that the label has adapted to digitization. This website may contain little information on the record company and it might not be used strategically.

We asked a second question concerning the distribution on digital music platforms. In France, in 2006, the three main platforms were the iTunes Music Store, VirginMega and Fnacmusic. Half of the labels were not available on these platforms, whereas almost a third (32%) were available on all three platforms. We also asked the labels in which year their music became available on digital music platforms. Roughly half of the labels reported that they started to be

[8] See, for example, Godes and Mayzlin (2004) and Chevalier and Mayzlin (2006).

distributed on digital platforms in 2005 and later, that is, relatively late compared to the downturn in music sales, which occurred between 2002 and 2003 in France. This suggests that the presence on digital platforms might be related to factors other than the adaptation to digitization.[9]

Hence, we focus on three variables that seem more relevant to capture adaptation to digitization. First, we consider that record companies have adapted to digital distribution if their current contracts with artists feature clauses on digital music distribution (*DIG-DISTR*). 81% of the labels reported that they included such specific clauses in their new contracts. Second, we specify as adaptation to digital scouting a dummy variable, which states whether the label has already recruited an artist through the Internet (*DIG-A&R*).[10] That was the case of 25% of the record companies in our sample. We also asked the labels whether they were currently using the Internet to recruit new artists. 66% of the labels never used the Internet in 2006 to find new artists, 24% used it "sometimes", and only 9% used it "often" or "always".

Finally, we consider that those labels which allow their artists to provide free streaming music on their websites are adapted to digital promotion the labels. This is measured by the dummy variable *DIG-PROMO*. Most labels (81%) authorized their artists to offer free streaming music, and many labels (49%) also authorized free downloads. Interestingly enough, even some for-profit labels authorized free downloads.

These three variables pertaining to adaptation to digitization are, of course, correlated to some extent. Labels that have already recruited artists on the Internet are more likely to authorize free music streaming (significant at 10%). They are also more likely to incorporate digital diffusion clauses in their current contracts (significant at 1%). Finally, labels that do not feature digital diffusion clauses in their current contracts are less likely to allow their artists to offer free music streaming (significant at 1%).

[9] In particular, we find that for-profit labels are available on 1.5 platforms on average, compared to 0.5 for not-for-profit labels (the difference is significant at less than 1%). Classical and jazz labels are available on 0.5 platforms against 1.5 platforms for labels from other genres of music (the difference is also significant at less than 1%). In other words, the distribution on digital platforms seems strongly related to the objective function of the label, and its main genre of music.

[10] This could be either because the artist had sent a digital demo file by email to the label, or because the record company had discovered the artist on a specialized web site like MySpace. Note that the name of our variable, "A&R", stands for "Artist&Repertoirs". This is how the scouting activity is usually referred to by record companies.

Therefore, though *DIG-DISTR*, *DIG-A&R* and *DIG-PROMO* correspond to different aspects of the adaptation to digitization, we also construct a composite variable, *ADAPT*, to distinguish between record companies that have adapted and those that have not. We specify the variable *ADAPT* as a dummy, which takes the value 1 if at least two of the three dummies *DIG-DISTR*, *DIG-A&R* and *DIG-PROMO* take the value 1, and 0 otherwise. According to this composite variable, *ADAPT*, 72% of the record companies in our sample have adapted to digitization, whereas 28% have not.

5. Results

In this section we first present our main estimation results, and then provide some robustness checks.

5.1 Results

In what follows, we begin by analyzing our first hypothesis, which states that adaptation to digitization does not increase the sales of a record company. We then proceed with our second hypothesis, which states that music labels that have adapted to digitization release more new albums than do labels that have not.

We begin by estimating equation (3) for the commercial output. The estimation results are provided in Table 1. Column (1) presents the results from the OLS estimation without adaptation variables, whereas column (2) introduces our three adaptation variables as explanatory variables, and column (3) uses the composite adaptation variable (*ADAPT*).

We start the discussion of the results with the input variables. In columns (1) to (3), input variables have the expected positive sign, and most of them are statistically significant. The two labor inputs, *EMPLOYEES* and *PR%*, have a significant and positive effect on sales. The capital input *CATALOG* also has a positive and significant effect on sales, as expected, whereas we find no significant effect for the number of *DEMOS* received each month. Finally, not surprisingly, for-profit record companies sell more CDs than non-profit labels.

Table 1 – Commercial output - LOG(*SALES*)

	(1)	(2)	(3)
LOG(*EMPLOYEES*)[11]	0.800***	0.811***	0.813***
	(0.236)	(0.230)	(0.235)
PR%	0.0133***	0.0132***	0.0131**
	(0.00420)	(0.00477)	(0.00475)
LOG(*CATALOG*)	0.386***	0.446**	0.454**
	(0.146)	(0.153)	(0.146)
LOG(*DEMOS*)	0.228	0.231	0.226
	(0.158)	(0.202)	(0.201)
FORPROFIT	1.262***	1.075**	1.187*
	(0.416)	(0.439)	(0.435)
MAJORDISTR	0.733	0.498	0.716
	(0.453)	(0.578)	(0.535)
COMPIL	0.483	0.324	0.355
	(0.386)	(0.398)	(0.390)
POPMUSIC	0.655	0.795	0.765
	(0.553)	(0.648)	(0.629)
EUROPE	0.393	0.308	0.290
	(0.453)	(0.461)	(0.465)
YOUNG	0.632	0.425	0.505
	(0.415)	(0.431)	(0.426)
AGEWEB	-0.129	-0.178**	-0.166***
	(0.0817)	(0.0816)	(0.0817)
DIG-A&R		-0.336	
		(0.447)	
DIG-DISTR		0.944	
		(0.576)	
DIG-PROMO		0.00310	
		(0.434)	
ADAPT			0.511
			(0.436)
Constant	3.546***	3.253***	3.359***
	(0.749)	(0.815)	(0.794)
Observations	105	98	98
R^2	0.599	0.625	0.617
Adjusted R^2	0.551	0.562	0.562

Heteroskedasticity-robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In columns (2) and (3), we introduce the three adaptation variables to test Hypothesis 1. We observe no major change in the sign or in the significance of the coefficients of the previous variables. The only difference is that the coefficient of the *AGEWEB* variable, which is negative, becomes significant at the 5% level. This suggests that the record companies that created their Web site most recently (i.e., with low values of *AGEWEB*) sell more CDs, ceteris paribus. This might be due to the fact that in markets where consumer tastes are changing fast, last movers benefit from information spillovers from first movers, and provide music that better suits consumers' tastes.

[11] The log(*EMPLOYEES*) variable is actually log(*EMPLOYEES*+1) since there are many 0s.

Column (2) shows that neither of the three adaptation variables is significant at the 10% level. We conducted a test for the joint significance on the three adaptation variables, which yielded F=0.95 and a p-value of 0.4208. We therefore fail to reject the null hypothesis (H0) that the coefficients of the three adaptation variables are all equal to zero, at the 1% level. Finally, the variable *ADAPT* in column (3) also allows us to test Hypothesis 1. We fail to reject the null hypothesis that the coefficient of *ADAPT* equals zero, and therefore, Hypothesis 1 is validated. To sum up, for the commercial output, we do not find any significant effect of the adaptation to digitization on the performance of record labels. Labels that have adapted to digitization do not sell more than labels that have not embraced the digital technology. Hypothesis 1 is thus supported by our empirical estimation.

Table 2 provides the estimation results for the creative output. Columns (1)-(3) present the results with the standard Poisson regression model. Column (1) gives the estimation results without adaptation variables, and columns (2) and (3) provide the results with the three adaptation variables and the composite variable (*ADAPT*), respectively. In all regressions, we add a lagged variable for the number of albums, *ALBUMS04*, to control for unobserved factors (such as, the efficiency of the label, or its creativity).

The effect of labor inputs is positive, and statistically significant for the press relations variable (*PR%*). The size of the catalogue has a positive effect, but is not significant, while the number of demo tapes received each month has a statistically significant and positive effect. We also find that for-profit labels produce more new albums than not-for-profit labels. This is consistent with the idea that for-profit labels have stronger incentives to release new albums to make sales, while not-for-profit labels may be limited by their financial capacities. Finally, record companies that frequently have titles in compilations (*COMPIL*) and that focus on mainstream popular music (*POPMUSIC*) produce fewer albums than do the others. This seems consistent with the logic of the star system (Burnett, 1996).

Table 2 – Creative output (*ALBUMS*)

	Poisson regression model			Poisson QMLE	Negative binomial
	(1)	(2)	(3)	(4)	(5)
ALBUMS04	0.0426***	0.0391***	0.0371***	0.0391***	0.0684***
	(0.00400)	(0.00594)	(0.00599)	(0.00943)	(0.0119)
EMPLOYEES	0.00115	0.00146	0.00153	0.00146	-0.000608
	(0.00111)	(0.00117)	(0.00115)	(0.00133)	(0.00264)
PR%	0.00547***	0.00586***	0.00546***	0.00586***	0.00499***
	(0.00103)	(0.00110)	(0.00108)	(0.00173)	(0.00166)
CATALOG	0.0000155	0.000602	0.000581	0.000602	-0.000559
	(0.000418)	(0.000625)	(0.000630)	(0.000977)	(0.00110)
DEMOS	0.00140**	0.00110*	0.00157***	0.00110	-0.000246
	(0.000558)	(0.000626)	(0.000607)	(0.000883)	(0.00126)
FORPROFIT	0.539***	0.338**	0.448***	0.338*	0.247
	(0.145)	(0.156)	(0.155)	(0.183)	(0.197)
MAJORDISTR	0.158	0.412***	0.210	0.412**	0.128
	(0.118)	(0.140)	(0.130)	(0.179)	(0.243)
COMPIL	-0.148	-0.204**	-0.179*	-0.204	-0.235
	(0.0979)	(0.0991)	(0.101)	(0.142)	(0.146)
POPMUSIC	-0.459***	-0.506***	-0.443***	-0.506***	-0.0922
	(0.113)	(0.117)	(0.115)	(0.184)	(0.188)
EUROPE	0.0323	0.102	0.0662	0.102	0.193
	(0.117)	(0.120)	(0.119)	(0.184)	(0.178)
YOUNG	0.361***	0.349***	0.378***	0.349*	0.189
	(0.0950)	(0.0998)	(0.0979)	(0.184)	(0.149)
AGEWEB	0.0166	0.0301	0.0114	0.0301	0.0357
	(0.0183)	(0.0193)	(0.0187)	(0.0344)	(0.0282)
DIG-A&R		0.465***		0.465**	0.101
		(0.0967)		(0.209)	(0.156)
DIG-DISTR		0.0374		0.0374	0.0479
		(0.171)		(0.293)	(0.218)
DIG-PROMO		0.372**		0.372*	0.643***
		(0.155)		(0.196)	(0.235)
ADPAT			0.277**		
			(0.123)		
Constant	0.719***	0.285	0.571***	0.285	-0.184
	(0.193)	(0.253)	(0.219)	(0.378)	(0.357)
α					0.168***
					(0.046)
Observations	106	99	99	99	99
Pseudo R^2	0.6170	0.6456	0.6260	0.6456	0.2418
Log likelihood	-297.00	-264.44	-279.08	-264.44	-230.26
LR χ^2	956.98	963.35	934.08	722.49	146.83

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In contrast with the commercial output, we find a strong and significant effect of adaption to digitization on the creative output. In particular, the *DIG-A&R* and *DIG-PROMO* variables have a positive and significant effect on the number of new releases. Moreover, the three variables of adaptation to digitization are jointly significant. A test of joint exclusion on *DIG-DISTR*, *DIG-A&R* and *DIG-PROMO* on model (2) yields χ^2=34.57, which corresponds to a *p*-

value of 0.0000. Hence, we reject the null hypothesis (H0') that the three coefficients are equal to zero at the 1% level. Similarly, in model (3), the composite adaptation variable (*ADAPT*) is statistically significant, with a positive sign. These results are consistent with Hypothesis 2. It suggests that adaptation to digitization substantially improves the creative efficiency of record companies, and that music labels that have adapted to digitization therefore release more albums.

A strong assumption of the Poisson regression model is that it assumes equi-dispersion, that is, that the (conditional) variance equals the (conditional) mean. To test the null hypothesis of equi-dispersion against the alternative of over-dispersion, we suppose that $\text{Var}(y|\mathbf{x})=\text{E}(y|\mathbf{x})+\alpha^2 \text{E}(y|\mathbf{x})$, where y is the dependent variables and $\mathbf{x}$ the vector of independent variables, and we test H0: $\alpha=0$ against H1: $\alpha\neq0$. Using the methodology proposed by Cameron and Trivedi (2009), we estimate α to be 0.094 with a standard error of 0.022, and we reject H0 at the 1% level. This result gives evidence of some degree of over-dispersion.

To account for over-dispersion, one solution is to estimate the Poisson regression model with the quasi-maximum likelihood estimator, and to use robust standard errors (see Wooldridge, 2002; Cameron and Trivedi, 2009). The estimation results are provided in Column (4) of Table 2. Our main result, that labels that have adapted to digitization release more new albums, still holds.[12]

Finally, we also use the negative binomial model, which is consistent with over-dispersion. The estimation results are provided in column (5) of Table 2. Note that the negative binomial estimate of the over-dispersion parameter (α) gives 0.168, which is close to our previous estimate of 0.094. A test of joint exclusion on *DIG-DISTR*, *DIG-A&R* and *DIG-PROMO* gives $\chi^2=9.26$, which corresponds to a *p*-value of 0.026. Therefore, we also reject the null hypothesis that the coefficients of the adaptation variables are all equal to zero at the 5% level.

[12] We reject the null hypothesis that the coefficients of the three adaptation variables are all equal to zero, at the 1% level.

To summarize, our results suggest that, in line with the long tail theory, adaptation to digitization makes record companies "sell less of more" music albums. Digitization enhances the creativity of record companies, leading digitized music labels to release more new albums. However, this does not result in higher sales for those labels. Probably, as Anderson (2006) posits, this is because the new releases of the digitized labels target niche markets with few consumers.

5.2 Robustness checks

In this section we provide two additional robustness checks, by splitting the sample for small and large labels, and for-profit and non-profit labels.

Table 3 in Appendix C provides the estimation results for the labels' total sales. Column (1) corresponds to the subsample of "small" labels (with 0 or 1 employee) and column (2) corresponds to the subsample of "large" labels (2 or more employees). For both small and large labels, we do not find any significant effect of adaptation to digitization. Splitting the sample between for-profit and not-for-profit labels (columns (4) and (4) in Table 3, respectively) yields the same result.

Table 4 (in Appendix C) provides the same robustness checks for the creative output (*ALBUMS*), for the QMLE Poisson regressions.[13] Columns (1) and (2) of Table 4 provide the results for small labels and large labels, respectively. We find that, whereas digitization has a positive impact on the creativity of large labels, it seems to have a negative impact on the creativity of small labels. Indeed, the sign of *DIG-A&R* is negative and significant for the smallest labels. One interpretation could be that, because artist selection becomes more efficient with digital tools, it leads small labels to become more selective. Columns (3) and (4) give the estimation results for for-profit, and not-for-profit labels, respectively. We find a positive and significant effect of digitization on creativity only for the for-profit labels.[14] These results suggest that digitization has a positive effect only for for-profit "large" labels.

[13] The same robustness checks for the negative binomial regressions yield the same results, and are available upon request from the authors.

[14] For the non-profit labels, a test of joint exclusion yields χ^2=2.24, and *p*-value=0.5234. Hence, we do not reject the null hypothesis that the coefficients of the three adaptation variables are all equal to zero.

6. Conclusion

In this paper we use data from a survey on 151 French record companies to test the prediction of the long tail theory (Anderson, 2006) that the future of the music business, in the digital era, is to sell fewer units of a larger number of creations (albums). We consider different aspects of adaptation to digitization: the digitization of artist scouting, the digitization of distribution, and the digitization of promotion. We use a production function approach, with two different types of output: commercial output (albums sales) and creative output (number of new albums released).

Our results suggest that, in line with the long tail theory, adaptation to digitization has a strong and positive impact on the production of new albums (the creative output), but no effect on sales (the commercial output). Digitization allows record companies to expand their catalogue by releasing more new albums that do however target niche markets. Selling fewer units of a greater number of albums does not have any positive impact on the overall sales of record companies.

References

Alexander P.J., (1994). New Technology and Market Structure: Evidence from the Music Recording Industry. *Journal of Cultural Economics*, 18(1): 113-123.

Anderson C., (2006). *The Long Tail*. London: Random House.

Belleflamme, P. and M. Peitz, (2010). Digital Piracy: Theory. CESifo Working Paper Series No. 3222.

Benghozi P.J. and F. Benhamou, (2010). The Long Tail: Myth or Reality? *International Journal of Art Management*, 12, 43–53.

Bishop P. and S. Brand, (2003). The Efficiency of Museums: a Stochastic Frontier Production Function Approach. *Applied Economics*, 35(17): 1853–1858.

Brynjolfsson, E., Hu, Y. J. and D. Simester, (2011). Goodbye Pareto Principle, Hello Long Tail: the Effect of Search Costs on the Concentration of Product Sales. *Management Science*, 57(8): 1373-1386.

Brynjolfsson, E., Hu, Y. J. and M.D. Smith, (2010). The Longer Tail: the Changing Shape of Amazon's Sales Distribution Curve, Working Paper. Available at http:// ssrn.com/abstract=1679991.

Burnett R., (1996). *The Global Jukebox*. London: Routledge.

Cameron A.C. and P.K. Trivedi, (2009). *Microeconometrics Using Stata*. StataCorp LP, College Station, Texas.

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

Easley R.F., Michel J.G. and S. Devaraj, (2003). The MP3 Open Standard and the Music Industry's Response to Internet Piracy. *Communications of the ACM*, 46(11): 91-96.

Elberse A., (2008). Should you Invest in the Long Tail? *Harvard Business Review*, 86(7/8): 88-96.

Elberse A. and F. Oberholzer-Gee, (2008). Superstars and Underdogs: An Examination of the Long-Tail Phenomenon in Video Sales. Mimeo.

Gapinski J.H., (1980). The Production of Culture. *Review of Economics and Statistics*, 62(4): 578–586.

Gapinski J.H., (1984). The Economics of Performing Shakespeare. *American Economic Review*, 74(3): 458–466.

Godes D. and D. Mayzlin, (2004). Using Online Conversations to Study Word of Mouth Communication. *Marketing Science*, 23 (4): 545-560.

Kumar A., Smith M.D. and R. Telang, (2011). Information discovery and the long Tail of motion picture content. Working Paper, Carnegie Mellon University, Pittsburgh, PA.

Liebowitz S., (2008). Testing File-Sharing's Impact by Examining Record Sales in Cities. *Management Science*, 54: 852-859.

Oberholzer-Gee F. and K. Strumpf, (2007). The Effect of File Sharing on Record Sales: An Empirical Analysis. *Journal of Political Economy*, 115(1): 1-42.

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

Peltier S. and F. Moreau (2012). Internet and the ''Long Tail versus Superstar Effect'' Debate: Evidence from the French Book Market. *Applied Economics Letters*, 19(8), 711-715.

Peterson R. and D. Berger, (1975). Cycles in Symbolic Production: The Case of Popular Music. *American Sociological Review*, 40: 158–173.

Premkumar G., (2003). Alternative Distribution Strategies for Digital Music. *Communications of the ACM*, 49(6): 89-95.

Regner T., (2003). Innovation of Music. In: R. Watt (ed), *The Economics of Copyright: Developments in Research and Analysis*, Cheltenham: Edward Elgar.

Taalas M., (1997). Generalized Cost Functions of Producers of Performing Arts – Allocative Inefficiencies and Scale Economies in Theatres. *Journal of Cultural Economics*, 21: 335-349.

Throsby D., (1977). Production and Cost Relationships in the Supply of Performing Arts Services. In: K. Tucker (ed.), *Economics of the Australian Services Sector*, London: Groom Helm.

Throsby D., (2006). An Artistic Production Function: Theory and an Application to Australian Visual Artists. *Journal of Cultural Economics*, 30: 1–14.

Wooldridge J.M., (2002). *Econometric Analysis of Cross Section and Panel Data*. The MIT Press, Cambridge, Ma.

Appendix A: List of variables

Variable	Description
SALES	Number of CDs sold in 2005
ALBUMS	Number of new albums produced in 2005
ALBUMS04	Number of new albums produced in 2004
EMPLOYEES	Number of employees in 2005
PR%	Percentage of albums that benefited from press relations in 2005
CATALOG	Number of music rights in the active catalog
DEMOS	Number of demo tapes received per month
POPMUSIC	Dummy variable indicating whether the label's main genre of music is popular music (pop, rock, soul, funk, hip hop, electronic music, techno, world, reggae)
FORPROFIT	Dummy variable indicating whether the label is a for-profit organization
MAJORDISTR	Dummy variable indicating whether the label is (even partially) distributed by a major
COMPIL	Dummy variable indicating whether the label had titles in a compilation in 2005
EUROPE	Dummy variable indicating whether the label is distributed in other European countries outside France
YOUNG	Dummy variable indicating whether the label was founded after 1999 (i.e., after the Internet bubble burst)
AGEWEB	Age of the website (equals 2006-year founded)
DIG-DISTR	Dummy variable indicating whether current contracts with artists take into account digital music distribution
DIG-A&R	Dummy variable indicating whether the label has already recruited an artist through the Internet
DIG-PROMO	Dummy variable indicating whether the label allows its artists to provide their music for free as streaming
ADAPT	Dummy variable indicating whether the label has adapted in at least two different respects: distribution, scouting, and promotion

Appendix B: Summary statistics

Table B1: Summary statistics (*N*=151)

Variable	Number of observations	Mean	Standard deviation	Minimum	Maximum
SALES	134	197,053	1,468,395	1.15	1.66e+07
ALBUMS	147	6.3	10.8	0	80
ALBUMS04	139	6.2	11.4	0	90
EMPLOYEES	145	6.3	22.4	0	200
PR%	139	33.9	43.4	0	100
CATALOG	144	58.8	133.5	0	800
DEMOS	147	30.4	56.9	0	500
POPMUSIC	149	0.76	0.43	0	1
FORPROFIT	151	0.70	0.46	0	1
MAJORDISTR	149	0.10	0.30	0	1
COMPIL	146	0.57	0.50	0	1
EUROPE	148	0.77	0.42	0	1
YOUNG	151	0.48	0.50	0	1
AGEWEB	136	4.5	2.7	0	10
DIG-A&R	149	0.25	0.43	0	1
DIG-DISTR	145	0.81	0.40	0	1
DIG-PROMO	139	0.81	0.39	0	1
ADAPT	135	0.72	0.45	0	1

Note: The descriptive statistics above use the full set of valid observations for each variable. The number of observations varies across variables due to missing answers.

Appendix C: Robustness checks

Table 3 – Robustness checks for the commercial output - LOG(*SALES*)

	(1) Small labels (0-1 employee)	(2) Large labels (2+ employees)	(3) For-profit labels	(4) Non-profit labels
LOG(EMPLOYEES)	0.229	1.103***	1.176***	-0.385
	(0.694)	(0.299)	(0.246)	(0.365)
PR%	0.0115	0.00926	0.0130**	0.0178
	(0.00766)	(0.00632)	(0.00594)	(0.0114)
LOG(CATALOG)	0.685**	0.323*	0.422***	0.907*
	(0.318)	(0.190)	(0.158)	(0.461)
LOG(DEMOS)	0.119	0.388	-0.0150	0.697**
	(0.312)	(0.291)	(0.234)	(0.244)
FORPROFIT	0.271	1.833**		
	(0.638)	(0.747)		
DISTRIBMAJOR	2.058	0.433	0.528	
	(1.219)	(0.728)	(0.644)	
COMPIL	0.520	0.429	-0.0285	2.098**
	(0.740)	(0.488)	(0.501)	(0.735)
POPMUSIC	0.652	0.859	0.847	1.426*
	(1.071)	(0.796)	(0.744)	(0.644)
EUROPE	0.918	0.140	-0.116	1.905**
	(0.649)	(0.686)	(0.661)	(0.718)
YOUNG	1.292	0.267	0.407	1.730
	(0.778)	(0.725)	(0.578)	(1.055)
AGEWEB	-0.0228	-0.249**	-0.221**	-0.456**
	(0.127)	(0.116)	(0.102)	(0.189)
DIG-A&R	-0.924	-0.215	-0.213	-0.564
	(0.753)	(0.563)	(0.521)	(0.696)
DIG-DISTR	0.991	0.803	0.260	-0.0160
	(0.636)	(0.934)	(0.704)	(0.785)
DIG-PROMO	-0.849	0.506	-0.230	-0.719
	(0.660)	(0.648)	(0.545)	(0.536)
Constant	3.006***	2.113	6.218***	1.588
	(0.946)	(1.388)	(1.412)	(1.093)
Observations	50	48	75	23
R^2	0.441	0.755	0.596	0.864
Adjusted R^2	0.217	0.651	0.510	0.701

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4 – Robustness checks for the creative output (*ALBUMS*)

	(1) Small labels (0-1 employee)	(2) Large labels (2+ employees)	(3) For-profit labels	(4) Non-profit labels
ALBUMS04	0.151***	0.0387***	0.0404***	0.365***
	(0.0348)	(0.00781)	(0.00933)	(0.139)
EMPLOYEES	0.327*	0.00210	0.00137	-0.0211
	(0.185)	(0.00163)	(0.00137)	(0.0282)
PR%	-0.000586	0.00444**	0.00597***	-0.00840
	(0.00239)	(0.00221)	(0.00189)	(0.00606)
CATALOG	-0.00334	0.000151	0.000488	0.00708
	(0.00355)	(0.000791)	(0.000962)	(0.0107)
DEMOS	0.0000692	0.0000499	0.00102	-0.00135
	(0.00343)	(0.000905)	(0.000879)	(0.0105)
FORPROFIT	-0.194	0.804***		
	(0.235)	(0.250)		
MAJORDISTR	0.752**	0.308*	0.409**	
	(0.357)	(0.161)	(0.183)	
COMPIL	-0.253	-0.0212	-0.181	-0.192
	(0.182)	(0.170)	(0.166)	(0.318)
POPMUSIC	-0.230	-0.481**	-0.532***	0.0717
	(0.222)	(0.190)	(0.192)	(0.329)
EUROPE	0.464	0.521**	0.115	0.0939
	(0.285)	(0.204)	(0.227)	(0.307)
YOUNG	0.377	0.566***	0.359*	-0.518
	(0.264)	(0.178)	(0.199)	(0.476)
AGEWEB	0.0330	0.00973	0.0337	-0.0696
	(0.0380)	(0.0408)	(0.0397)	(0.0694)
DIG-A&R	-0.531**	0.527**	0.503**	-0.288
	(0.228)	(0.245)	(0.232)	(0.362)
DIG-DISTR	0.335	-0.0574	0.0562	0.123
	(0.225)	(0.286)	(0.378)	(0.282)
DIG-PROMO	0.417	0.268	0.303	0.400
	(0.318)	(0.238)	(0.216)	(0.359)
Constant	-0.378	-0.133	0.618	0.284
	(0.741)	(0.483)	(0.484)	(0.416)
Observations	50	49	75	24
Pseudo R^2	0.5490	0.7156	0.6499	0.2482
Log likelihood	-88.59	-131.05	-220.25	-35.15
LR χ^2	1026.47	853.89	498.65	197.33