European Commission

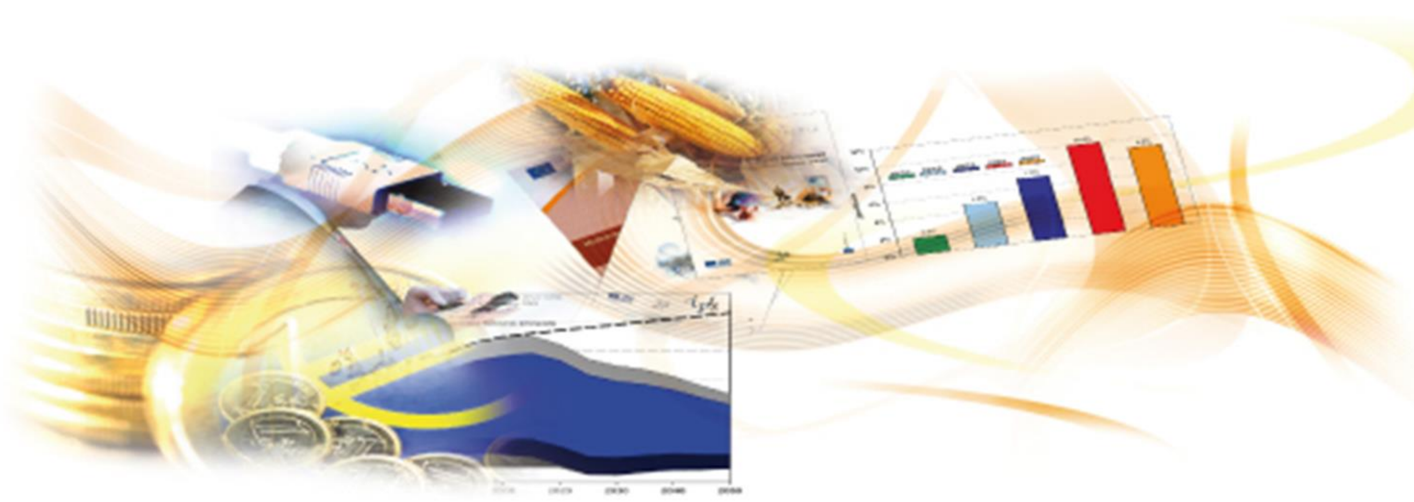
JRC TECHNICAL REPORTS

Digital Economy Working Paper 2014/09

Search Costs, Information Exchange and Sales Concentration in the Digital Music Industry

Néstor Duch-Brown
Bertin Martens

2014



Joint Research Centre

European Commission
Joint Research Centre
Institute for Prospective Technological Studies

Contact information
Néstor Duch-Brown
Address: Edificio Expo. c/ Inca Garcilaso, 3. E-41092 Seville (Spain)
E-mail: jrc-ipts-secretariat@ec.europa.eu
Tel.: +34 954 488 318
Fax: +34 954 488 300

https://ec.europa.eu/jrc
https://ec.europa.eu/jrc/en/institutes/ipts

This publication is a Technical Report by the Joint Research Centre of the European Commission.

JRC90511

ISSN 1831-9408 (online)

Spain: European Commission, Joint Research Centre, 2014

Abstract:

It is often assumed that consumers benefit from the internet because it offers a “long tail” with more variety of products to choose from. However, search costs may block the long tail effect and result in the dominance of superstars. This paper examines the variety hypothesis in the entire online market for digital music downloads in 17 countries over the period 2006-2011. First, we show that the entire distribution of legal music downloads is heavily skewed. Second, we hypothesise that a wide range of online information channels (sales and discovery platforms) play a role in this market. We find that the reduction of search costs implied by the generalisation of online information tools transforms demand as a result of changes in the dispersion of preferences. Ubiquitous and very popular discovery channels such as Facebook and iTunes tend to push consumers towards the superstars by shifting the demand curve but also towards the long-tail since they also generate rotations that promote niches. Consequently, both the superstar and the long tail effects emerge even in mature digital markets.

Keywords: Digital markets, search costs, information, sales concentration

JEL-codes: C46; D12; L82

1. Introduction[1]

Consumers are assumed to benefit from the internet in several ways. The internet vastly reduces the cost of communication and thereby increases the supply of information on goods and services that consumers may wish to buy. This increases the variety of goods known to consumers and available for consumption. It may also enhance price competition and economies of scale in online product markets as suppliers can now reach a geographically much wider market and consumers can select from a much larger number of geographically spread suppliers. These potential benefits have been well-studied. A key assumption for the realisation of these benefits is that consumers can effectively handle the increased supply of information to make more informed choices. This raises the question of search costs. While the supply of information available to the consumer has vastly increased, we can safely assume that his internal cognitive processing capacity has not significantly changed following the rise of the internet and digital technologies. Consumers can use external information processing tools to pre-select among the available options. Search engines for instance are such an external pre-processing tool: they rank available options in accordance with a designed algorithm that do, to some unknown extent, take into account the user's implicit and explicit preferences. Online price comparison engines or recommendation systems are yet other examples.

The objective of this paper is to empirically investigate to what extent consumers manage to screen the huge amount of information that the internet provides them with and find what they are searching for with the help of information exchange, through word of mouth or external information intermediaries and processing tools. The question is whether these external tools guide the consumer to his own preferred choices or to some "average" choices of a wider set of consumers. In other words, what is the trade-off that these information processing intermediaries induce between conformity and diversity? In internet economics jargon, this translates into a trade-off between superstars (Rosen, 1981) and the long tail (Anderson, 2006).

There has been a fair amount of empirical research on this question over the last decade. However, the findings remain very much divided between studies that show the dominance of the long-tail phenomenon and those that find evidence in support of the superstar effect. A characteristic of all these empirical studies is that they focus on a specific company or online sales platform such as Amazon books (Brynjolfsson et al., 2006) or a

[1] We thank participants at the IPTS Digital Economy seminar and at the 7th ICT Conference Paris on The Economics of Information and Communications Technologies for useful comments and suggestions. In addition, particular thanks go to Joel Waldfogel, Michail Batikas, Thomas Frick, Dimitrios Tsekouras, Frank Verboven, Lukasz Grzybowski, Lisa George and François Moreau for insightful discussions and suggestions. However, we bear all the responsibility for any remaining errors or omissions.

clothing sales company (Brynjolfsson et al., 2011). In this paper, we examine consumer behaviour in a specific product market –à la carte digital music downloads in Europe and North America- where a variety of suppliers, online platforms and information intermediaries operate. All these market participants may have their own information biases that they transmit to consumers; however, it is the aggregate effect that we are looking for. Moreover, we observe this market in different stages of maturity across a range of countries and years.

Between 2006 and 2012 the number of global music downloads, a combination of single tracks and digital albums, went from 1.2 billion to 4.7 billion (IFPI, 2013). We use data on digital music downloads by song title in 17 countries in Europe and North America over the period 2006-2011. Our data shows that the number of distinct songs available online in the countries in our dataset increased from 1.7 to 4.7 million between 2006 and 2011. Clearly, online music stores offer a much wider variety of supply than traditional bricks and mortar record stores could ever provide.[2] At the same time, walking around and browsing in an online store is much more difficult for consumers. This is where information intermediaries such as search and recommender systems –as well as other forms of online information exchange such as social networks- play an increasingly important role. The question is how digital information and distribution technology has affected this information stream and decision making process.

Our results show that the distribution of legal music downloads is heavily skewed. This implies that sales concentration is high. For instance, on aggregate, the world's top 50 songs in the period 2006-2011 represent 35% of world sales. Taking into account that we observe more than six million songs, this share is highly significant. Moreover, we also observe that around 30% of songs registered only one download in the entire period of six years. In addition, we also offer evidence supporting the idea that the reduction in search costs implied by the generalisation of online information tools transforms demand as a result of changes in the dispersion of preferences. Ubiquitous and very popular discovery channels such as Facebook and iTunes tend to push consumers towards the superstars by shifting the demand curve but also towards the long-tail since they also generate rotations that promote niches. Consequently, both the superstar and the long tail effects emerge even in mature digital markets.

The paper is organised as follows. In Section 2, we briefly review the related literature on the superstar and long-tail effects. In Section 3 we describe the data and we show some of the characteristics of legal music downloads distribution. Section 4 explains the

[2] Several estimations suggest this figure to be between 40,000 and 50,000 songs or, equivalently, 3,000-4,000 albums (see Page and Garland, 2009).

methodology used to estimate the parameters of the distribution to assess the rank-downloads relationship. Section 5 discusses the results. In Section 6 we relate the evolution of sales concentration to the reduction in search costs implied by the penetration of online information tools. Finally, Section 7 offers some conclusions.

2 Literature review

The increasing use of digital technologies in cultural industries leads to a reduction in production costs, facilitating the appearance of more products into the market. Since distribution costs can be reduced as well, it becomes easier for these new products – particularly niche products- to enter the market. In order to help consumers to make appropriate choices, or to find those products that better match their preferences, digitisation has also allowed for the development of tools that ensure a better match between supply and demand in electronic markets. The combination of an enlarged variety of products along with a decrease in search costs increase the likelihood that more niche products are discovered and adopted. This in turn alters the traditional distribution of sales observed in offline markets, highly concentrated in a small number of superstars (Rosen, 1981; Adler, 1985). In online markets, the digital revolution will benefit less popular artists and will push the development of niche markets, a phenomenon defined as the long tail (Anderson, 2006; Brynjolfsson et al., 2006).

According to the original description of the superstar effect (Rosen, 1981), differences in talent (poor talent cannot substitute superior talent) and nearly perfect reproducibility of art generate stardom-like market outcomes. Here, the more talented the artist is the more advantages from operating in a large market since the more talented will be attracting massive earnings compared to the less talented. In addition, the outcome is efficient since a more talented artist can always enter and become a superstar. In contrast, Adler (1985) focused on the demand side, where artistic consumption is assumed to depend on obtaining and sharing information about a given artistic experience. In this setting, the increasing marginal utility of artistic consumption generates a learning process where search costs are minimised if consumers specialise in the most popular artist and by interacting with others, generates positive network externalities, even if there are no differences in talent.

Some authors, (Hamlen, 1991 and 1994; Chung and Cox, 1994; Strobl and Tucker, 2000; Crain and Tollison, 2002; Krueger, 2005), have conducted empirical tests of these conjectures for the traditional (physical) music industry, but research analysing the digital side of the music industry is scarce. However, evidence for other online markets is readily available. For example, Tucker and Zhang (2007) using an empirical comparison of consumers' click-through behaviour between the catalogue and Internet channels of a wedding services website, show that sales from superstar products are enhanced by attracting new demand without affecting the demand for niche products. Ghose and Gu (2007) find that consumer search costs for price information in two online book stores are lower for hit products compared to niche products which is consistent with the superstar effect.[3] With data from Netflix, the largest online movie rental company, Tan and Netessine (2009) find no evidence that niche titles satisfy consumer preferences better than hit titles.

The long tail phenomenon appeared when comparing the differences between offline and online retail sales under the potential effects of digitisation (Anderson, 2006; Brynjolfsson et al., 2006). Brynjolfsson et al. (2003) provide not only evidence of a significantly larger product variety at online book retailers, but also suggest that the Internet significantly lowers search costs allowing consumers to find more products. From these original contributions, the more recent literature has focused almost exclusively in electronic markets. Zhou and Duan (2012) argue that since the long tail explanation describes the change in consumption patterns it can only be examined comparing consumer demand distribution shifts over time in pure online channels. Using sales data from the craft beer industry and reviews from a Web-based rating service, Clemons et al. (2006) show that receiving the most positive reviews helps new products grow more rapidly in the marketplace, reducing sales concentration. Several recent studies have examined the long tail phenomenon in digital markets for various products: books (Oestreicher-Singer and Sundararajan, 2009; Brynjolfsson et al., 2010), clothing (Brynjolfsson et al, 2011), videos (Kumar et al., 2011), and software (Duan et al., 2009).

Digital music markets have been also scrutinised by the long tail lens. Using data from a leading music blog aggregator, Dewan and Ramaprasad (2012) analyse the relationship between music blogging and full-track sampling. Their findings suggest that the intensity of music sampling is positively associated with the popularity of a blog among previous consumers. Additionally, this association is stronger in the tail than in the body of music sales distribution. Bourreau et al. (2013) analyse the impact of digitisation on record companies. The authors argue that in line with the long tail theory, record labels that have adapted to digitisation should release more albums, selling less than the average album since their focus lies in filing niche markets. Their findings suggest that digitised record companies in France sell less (or at least not more) quantity of an increasing number of albums. These papers test whether the long tail phenomenon is emerging in digital music

3 A theoretical explanation supporting this finding states that online recommendation systems are biased towards more popular products since they are based on historical data and hence they help reduce sales diversity (Fleder and Hosanagar, 2009).

markets.[4] However, they use limited data sets to test their hypothesis. No research yet has tackled the issue of the complete distribution of legal digital music downloads. This paper aims to make contributions in that direction.

The long tail indicates a shift of demand from the hits to the niches, but the very popular products can still dominate market demand at the same time. Hence, the long tail and superstar phenomena are not necessarily conflicting and could coexist. For example, Elberse and Oberholzer-Gee (2007) found empirical evidence for the coexistence of the long tail and superstar phenomena by examining overall video sales through both online and offline channels. Zhou and Duan (2012) also demonstrate that the long-tail and the superstar effects can occur simultaneously. In their case, they suggest that product variety weakens the effect of online reviews and this reinforces the relative importance of extremely-ranked online software downloads.[5]

A key element supporting the long tail phenomenon is related to information transmission. Regardless of the platform used, the decision to buy music is shaped by personal preferences and social influences, both subject to information derived from peers, word-of-mouth and public sources (reviews, mass media or specialised magazines). In an early and influential contribution, Bakos (1997) show that online recommendations would help consumers find less popular goods that yet match their preferences. Since then, online product feedback and recommendation systems have been viewed as important tools to help reduce search costs in a context of ever increasing product variety. Since then, the literature on electronic word-of-mouth (eWOM) has exploded but the conclusions regarding its effects across products with different popularities are not consistent (Godes and Mayzlin, 2004; Chevalier and Mayzlin, 2006; Liu, 2006; Zhao et al., 2008; Duan et al., 2008a, 2008b and 2009; Goldmanis et al., 2009; Zhu and Zhang, 2010; Chen and Xie, 2011).[6] As a result, there is a growing literature devoted to the analysis of product search and search costs in computer-mediated environments (De los Santos et al. 2012; Baye et al., 2013a and 2013b).

Building on the contribution of Johnson and Myatt (2006), Bar-Isaac et al. (2012) address theoretically the relationship between consumers' search costs, information transmission and sales concentration. The authors interpret the diffusion of search engines, the Internet,

4 Other authors have dealt with the issue of music production diversity and examine whether we can observe a trend toward the homogenization of musical production and a negative impact of market concentration on diversity. In particular, see Peterson and Berger (1975, 1996); Burnett (1992); Lopes (1992); Alexander (1994, 1996) and Dowd (2004).

5 These papers offer evidence consistent with some theoretical contributions that allow for this coexistence. See, for example, Kendall and Tsui (2011) and Weeds (2012).

6 Theoretical contributions have also been elaborated around this issue. See, for instance, Fleder and Hosanagar (2009), Hervas-Drane (2013) and Bar-Isaac et al. (2012).

and communication and information technologies more generally as a fall in search costs. Their analysis leads naturally to long-tail and superstar effects arising simultaneously. In particular, they focus on how lowering search costs can have remarkable effects on sales distributions. When search costs go down, consumers will search longer and the likelihood of finding high quality firms is higher. In this case, better firms are more successful than before, leading to a superstar effect. In addition, consumers are also more likely to find products that best match their preferences inducing some firms to switch to niche designs, and hence creating a long-tail effect. Hence, the superstars and the long-tail both increase market share at the expense of firms in the middle of the distribution. These firms, facing a more competitive environment, choose niche designs with lower sales and higher mark-up, attracting additional buyers.

3 Data

Our dataset includes data on digital music downloads for 17 countries in Europe and North America for the period 2006-2011. Data comes from Nielsen and includes all songs downloaded from major online retailers such as iTunes, Amazon.com and others. Nielsen claims that this database covers around 80% of the world market for digital downloads. It contains 6.7 million different songs from 1.3 million different artists accounting for 7.6 billion downloads in the period under study. However some 75,000 artists account for 90% of world sales. Table 1 reports some descriptive statistics. As can be seen from the table, the dataset covers the most important markets for digital music and leading countries in internet penetration and e-commerce. Moreover, the data covers the entire active catalogue, i.e. songs with at least one observed download. The entire catalogue is likely to be larger but there is no information on never downloaded songs. Hence, we are able to analyse the entire active catalogue, not only the top-N upper-tail part of the distribution, as in most previous studies. Additionally, we can distinguish the distribution of downloads at the song level and at the artist level.

As the data in Table 1 shows, the distribution of downloads seems to be quite skewed: on aggregate, 30% of all songs have only one download in the entire period 2006-2011. However, the distributions by country differ from the aggregate since many songs that are only downloaded once per country are not reflected in this overall share. This is important since only a very small fraction of total songs have been downloaded in all 17 countries (less than 1%). For instance, in Finland, 64% of songs registered only one download while in the US the share was only 27%. For all countries in our data, the share of songs with only one download is significant. In addition, the share of songs with at least 1,000 downloads indicates a substantial skewedness of the distribution, ranging from a minimum of 0.03% in Finland to 1.6% in the US. Another characteristic of the distribution of downloads resides in the difference between the average downloads by song with their maximum values, reaching several orders of magnitude in every country. Although not shown, the figures for artists are qualitatively similar.

Table 1: Data description – songs

	Obs.	Downloads by song			Share of songs with	
		Min	**Avg.**	**Max**	**1 download**	**More than 1000 downloads**
Austria	2,392,627	1	8.1	35,829	57.9	0.06
Belgium	3,272,883	1	10.7	51,435	55.7	0.11
Canada	8,071,002	1	36.3	588,081	43.9	0.48
Switzerland	4,484,358	1	11.1	86,145	51.8	0.11
Germany	8,096,139	1	26.1	391,256	45.5	0.28
Denmark	2,477,841	1	14.3	67,612	55.7	0.16
Spain	2,722,095	1	12.4	173,188	57.2	0.10
Finland	1,183,042	1	4.8	8,929	63.6	0.03
France	6,267,545	1	21.7	230,125	48.0	0.27
United Kingdom	11,600,000	1	53.8	973,060	41.0	0.55
Ireland	1,955,614	1	12.3	47,439	57.4	0.13
Italy	3,708,194	1	13.7	134,664	54.4	0.16
Netherlands	3,059,281	1	7.9	51,029	58.2	0.07
Norway	2,590,004	1	10.8	83,083	55.5	0.11
Portugal	892,679	1	4.7	16,577	64.7	0.03
Sweden	2,582,481	1	7.7	26,757	59.2	0.07
United States	27,400,000	1	219.7	5,755,773	26.9	1.57

Since downloads can be interpreted as a measure of the song's popularity, we compute song rankings by country and year where the most downloaded song is given rank 1. Summary statistics for yearly downloads and yearly ranks are provided in Table 2. The heavily skewed distribution with large variance is again apparent. The calculated means are consistently and significantly larger than the medians. The mean of yearly downloads ranges from 92.6 in 2006 to 74.5 in 2011 while the median remains constant at 2. In terms of ranking, it can be observed that the average ranking of a given song is quite far away from the top ranked song and again far above the medians for all the years in the database.

Table 2: Song year downloads and year rank summary statistics

Downloads	mean	median	std. dev.	min	max
2006	92.6	2	3724.7	1	1,935,974
2007	89.7	2	4339.4	1	2,713,920
2008	90.8	2	4812.8	1	3,419,836
2009	83.4	2	5145.9	1	4,676,087
2010	74.7	2	5210.1	1	4,346,572
2011	74.5	2	5302.1	1	5,755,773
Ranks	**mean**	**median**	**std. dev.**	**highest**	**lowest**
2006	516,416.9	254,123	595,565	1	2,303,965
2007	746,331.0	381,156	834,210	1	3,322,808
2008	965,636.3	502,300	1,072,277	1	4,273,554
2009	1,186,742.0	644,743	1,310,788	1	5,271,065
2010	1,356,930.0	744,463	1,497,010	1	5,982,655
2011	1,423,570.0	808,081	1,554,785	1	6,271,424

These descriptive statistics point to an extremely skewed sales distribution. Figure 1 plot the distribution of songs' cumulative downloads in the period 2006-2011. Panels a) to c) in the figure partition the data into the top 500, top 5000, and top 50000 cumulative songs downloads while panel d) shows the entire distribution. The different panels in Figure 1 confirm visually the insights from the descriptive statistics of Tables 1 and 2: the distribution of digital music downloads is heavily skewed implying that sales concentration is high (the top 50 songs concentrate on average 35% of world sales) and the tail is quite long (2 out of 6.7 million songs had only one download in the period under analysis). Although we do not report the corresponding figures, this pattern is also repeated if we focus on specific countries, years and country-year pairs.

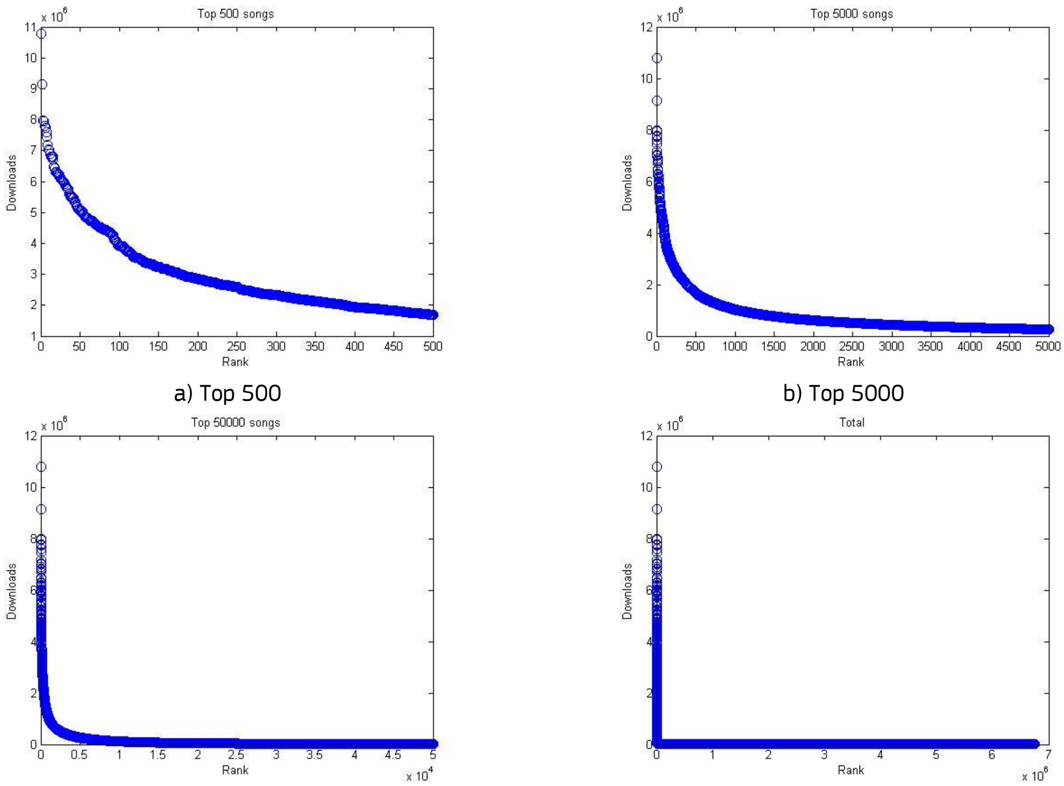
4 The distribution of music downloads

The aim of this section is to analyse the distribution of downloads. Figure 1 has confirmed that our measure of popularity (downloads) is characterised by extreme imbalances, as has been widely documented in the literature (Easley and Kleinberg, 2010). The probability density function for aggregate and cumulative downloads depicted in the figure looks like a Pareto distribution,[7] originally developed to describe the allocation of wealth among individuals (Pareto, 1897) but similar patterns have been observed in many other areas in

[7] The Pareto distribution belongs to the more general family of Power Law probability distributions which have special properties. A power law implies that the probability of occurrence of an event starts high and peters out. Few events occur very often while many others occur rarely. It follows that in a power law distribution the tails fall in accordance with an estimated power factor.

natural and social sciences.[8] A convenient feature of this type of distributions is that, in logarithms, the relationship turns out to be a straight line.[9]

Figure 1: Distribution of aggregated downloads of songs



a) Top 500 b) Top 5000

c) Top 50000 d) Total

Formally, the Pareto distribution[10] implies the following relationship between downloads and its rank:

$$DR^{\beta} = A \tag{1}$$

where A and β are constants, D denotes the number of downloads and R denotes the download rank. The natural logarithmic transformation of (1) yields the following equation

$$ln\ D = ln\ A - \ \beta\ ln\ R \tag{2}$$

---

8 The Pareto law appears to hold in many unconnected areas. See Clauset (xxx) and Newman (xx) for lists of applications in the natural and the social sciences.

9 The rank-size relation is extremely useful as a graphical device as it accentuates the tails of the empirical distribution. In particular, the appearance of a linear relationship in a double log scale can be interpreted as the sign of a power law distribution, which corresponds to the asymptotic limit of the Pareto distribution.

10 Simon (1955) showed that the Pareto law can be derived given two assumptions: i) Gibrat's law, stating that the growth rate of sales is size independent and; ii) a constant entry of new songs (artists).

meaning that on a log-log scale, β measures the slope of a straight line. Following some contributions of the literature (Brynjolfsson et al., 200x; Chevalier and Golsbee, 2003; Ghose and Gu, 2006; Hinz et al., 2011) we assume this log-linear relationship between sales of each product and its corresponding rank as a demand relationship. Hence, adding an error term to (2) we have

$$ln\, D = \alpha - \beta\, ln\, R + \varepsilon_{ict} \tag{3}$$

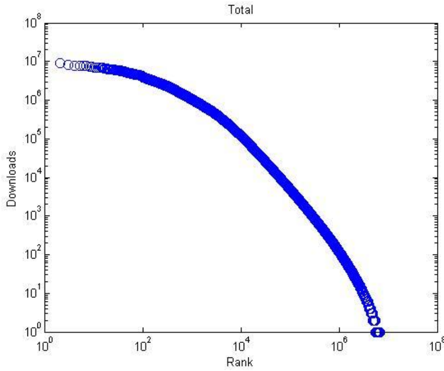
where $\alpha = lnA$ and ε is an independent and identically distributed random error. In this case, the parameter α indicates the intersection with the vertical axis and the parameter β measures the speed at which the demand per product falls as the rank increases.

However, visually, the logarithmic transformation of the distribution suggests a non-linear relationship between the variables. Figure 2 plots the rank-downloads relationship in log scale. In the figure, panel a) depicts the aggregated cumulative rank-downloads relationship (all countries and years aggregated). Panel b) shows the aggregated rank-downloads distribution in the initial (2006) and final (2011) years of our data. Panels c) and d) show the same relationship but for two separated samples of different countries included in our dataset. As the different panels of the figure show, the rank-downloads relationship in log scale is far from a straight line and shows a significant downward concavity. This means that the decay in downloads as rank increases is not as fast as a straight line would otherwise suggest.

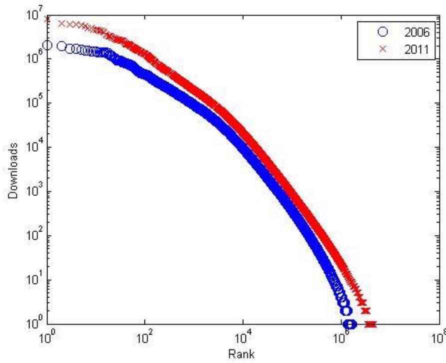
Two alternative explanations are in place. First, if we imagine a straight line connecting the two extremes of the rank-size relationship in panel a) of Figure 2, then we observe that middle-ranked products get more downloads than equation (3) would predict. Second, if we draw a straight line tangent to the midpoint of the distribution in panel a) of Figure 2, we observe that both superstars (i.e. the head of the distribution) and the long-tail get less downloads than equation (3) would predict as well. Ijiri and Simon (1974) observed that empirical rank-size distributions frequently deviate from log-linearity by exhibiting strong concavity. They suggested that the curvature of the distribution may be quantified by adding a quadratic term, as shown in the following equation:

$$ln\, D = \alpha - \beta\, ln\, R + \gamma\, (\ln R)^2 + \varepsilon_{ict} \tag{4}$$

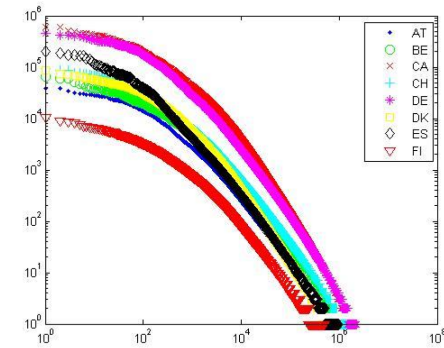
Figure 2: Distribution of aggregated songs' downloads



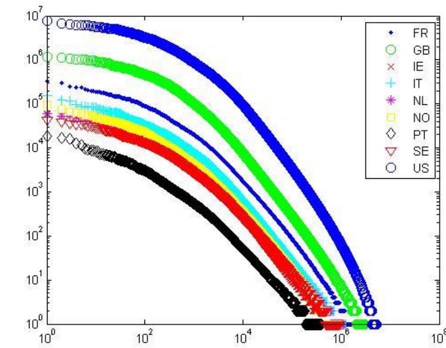
a) Total



b) Year (initial and final)



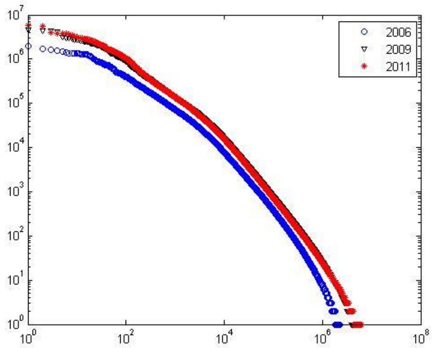
c) Countries (sample)



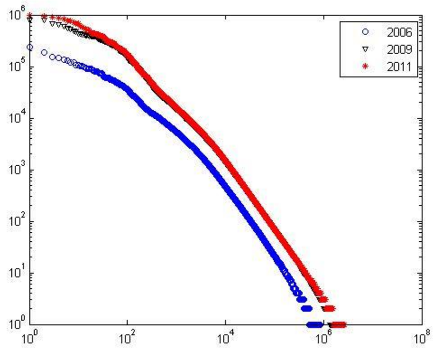
d) Countries (sample)

The curvature of the distribution is concave downward if the coefficient γ is negative and convex downward if it is positive. According to what we observe in Figures 2 and 3, we expect the estimated γ to be negative. In Section 2 we argued that, as search costs are reduced and information exchange increased due to the penetration of information and communication technologies, we should observe two different effects. On one hand, consumers search longer and share more information, increasing the probability of finding popular products, for which more information is available and thus exchanged. On the other hand, some consumers will increase their likelihood of finding products that best match their preferences, creating or expanding the long-tail. Hence, superstars and the long-tail arise simultaneously at the expense of middle-ranked products (Bar-Isaac et al., 2012).

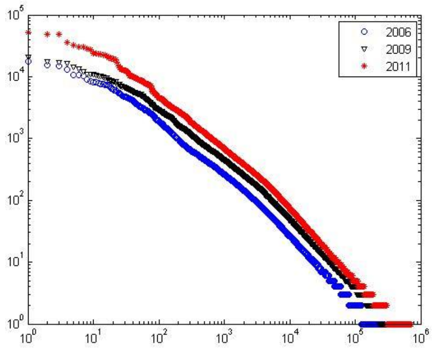
Figure 3: Evolution of songs' downloads in selected countries



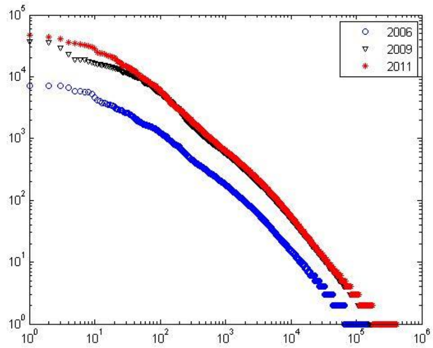
a) US



b) UK



c) Netherlands



d) Ireland

In order to test this argument, we should look at the evolution of the distribution of downloads. Panel b) in Figure 2 shows how the relationship between downloads and rank evolved on aggregate in the period 2006-2011. Additionally, Figure 3 also shows the evolution of this relationship for selected countries from the dataset. Panel a) plots the rank-downloads relationship for the US, the biggest country in terms of digital music donwloads representing 30% of all observations. The remaining panels show European countries of different sizes in terms of their digital music industry: UK in panel b), the biggest European digital music market; the Netherlands in panel c), an intermediate digital music industry and; Ireland in panel d) representing a small industry.

From these plots we get some insights on the evolution of the rank-downloads relationship. Looking for instance at the panel b) of Figure 2, we see that while the middle of the distribution shifts slightly to the right, the extremes –both the head and the tail- deviate significantly in 2011 with respect to its position in 2006. Hence, demand for digital downloads not only shift outwards -as the number of consumers increases most probably- but also rotates favouring both superstars and the long-tail. With small differences at the country level, this pattern is confirmed by Figure 3.

In the empirical ground, the introduction of a squared term into equation (4) could lead to problems of multi-collinearity, since ln(rank) and its squared term will be highly correlated. However, in the strictest sense, multi-collinearity refers only to linear relationships between variables. In any case, one approach to the potential problem of multi-collinearity in models with a polynomial term is to subtract the continuous variables from their means. When this was carried out on our data it produced no qualitative differences in the results. In addition, in order to correct for heterosckedasticity, the White procedure to compute standard errors was adopted in every regression. Additionally, we performed some robustness checks with robust estimation procedures that take into account the existence of outliers and also corrects for the presence of heteroskedasticity in the error term. We did not find any substantial modification from the OLS results reported in what follows.

5 Results

We estimate equation (4) using OLS and pay particular attention to the value of the gamma coefficient that indicates the speed at which downloads decrease with rank. We then try to explain variations in gamma in the next section. We concentrate on the results at the song level since the results at the artist level are similar. First, we concentrate in the entire universe of songs and perform different estimations with different aggregations of the data. In these regressions we introduce time and country dummies where appropriate. Then, we estimate the regressions year by year controlling for country specific effects with the objective of analysing the evolution of the parameter of interest (γ). Finally, we performed country level regressions controlling for year effects, with the intention to detect differences in the concavity of the rank-downloads relationship in the different countries of our sample.[11]

The results reported in Table 3 show a remarkable similarity between the regressions that use individual songs' downloads and the regressions that use different forms of cumulative downloads by time (specification 2), country (specification 3) and overall (specification 4). These results suggest that the rank-downloads relationship for songs shows significant concavity. The estimated gamma coefficient on the $[\ln(\text{Rank})]^2$ variable is negative and significant. Middle-ranked songs earn a disproportionately large share of downloads and this leads to a downwardly convex relation between rank and downloads.

[11] We performed alternative regressions with different sub-samples of the data. For instance, we replicated the whole exercise with the share of downloads instead of total number of downloads and the results were almost identical. We also restricted the sample to new songs only; to domestic songs; and to the top 75000 artist. In all these cases, the results were qualitatively unaltered.

Table 3: Estimates of the Rank-Downloads relationship

		Aggregation		
	(1)	(2)	(3)	(4)
ln(rank)	0.758***	1.620***	1.815***	3.356***
	(0.000940)	(0.00328)	(0.00473)	(0.0159)
$[\ln(\text{rank})]^2$	-0.0960***	-0.138***	-0.146***	-0.202***
	(3.75e-05)	(0.000129)	(0.000178)	(0.000572)
Constant	3.650***	2.267***	4.031***	-3.035***
	(0.00585)	(0.0207)	(0.0312)	(0.110)
Country dummies	Yes	Yes	-	-
Time dummies	Yes	-	Yes	-
Observations	92,919,358	25,899,550	20,274,656	6,752,059
R-squared	0.984	0.991	0.987	0.987

Note: Robust standard errors in parentheses. *** denote significant at 1%, ** at 5% cent and * at 10%, respectively.

Moreover, we find that the value of the gamma parameter is decreasing over time and with market size (see Table 4). This decline occurs in a period in which digital music downloads have been growing in importance, replacing traditional physical music sales as the main source for revenue for the industry. There are important differences by country. We find that all countries show negative and statistically significant estimates of the γ parameter, although with important differences in its value. For instance, the strongest effect is found in the largest market, the US (-0.128). As an additional consistency check, we aggregated song information by artist and did the same regressions at artist level. The results are not shown, but the γ parameter remained negative and significant in all regressions.

Table 4: Estimates of the Rank-Downloads relationship by year

	2006	2007	2008	2009	2010	2011
ln(rank)	1.213***	1.046***	0.984***	0.896***	0.746***	0.706***
	(0.00434)	(0.00334)	(0.00288)	(0.00257)	(0.00223)	(0.00214)
$[\ln(\text{rank})]^2$	-0.124***	-0.113***	-0.108***	-0.101***	-0.0934***	-0.0916***
	(0.000184)	(0.000138)	(0.000116)	(0.000102)	(8.78e-05)	(8.38e-05)
Constant	2.140***	3.184***	3.706***	4.154***	4.909***	5.220***
	(0.0253)	(0.0201)	(0.0177)	(0.0161)	(0.0141)	(0.0135)
Observations	7,213,178	11,234,039	14,425,149	18,146,582	20,241,147	21,659,263
R-squared	0.990	0.992	0.992	0.993	0.994	0.994

Note: all specifications include country dummies. Robust standard errors in parentheses. *** denote significant at 1%, ** at 5% and * at 10%, respectively.

These results are very similar to those obtained in previous studies, although very few report the evolution of the distribution of sales. Although the magnitude of our estimated γ is somewhat different from what we observe in the film industry – for instance, estimates in the literature range from -0.34 in Hand (2001) to -0.54 in the study of McKenzie (2008) - they are quite similar to the results obtained in the physical music industry by Giles (2007) who obtained an estimate of -0.11 and Elliott and Simmons (2011) who found a parameter of -0.11. An important difference between our study and others is the amount of information used. Here, we are using the whole distribution of songs, whereas previous analyses have only analysed a restricted upper part of the distribution.

6 Explaining demand dynamics in the digital music industry

The previous section showed a significant concavity in the size-downloads relationship of digital music downloads for 17 countries. An important implication of the observed curvature in the distribution is the relative dominance of middle-ranked observations. Both superstars and the long-tail are underrepresented. According to Bar-Isaac et al. (2012) this corresponds to a situation in which search costs are high enough to block many consumers and information transmission is limited.

We also reported a decrease over time in the value of the gamma coefficient and thus a decrease in the degree of concavity: demand is shifting outwards but also is rotating at the same time, increasingly looking more as a straight line. During the period under analysis, both the head and the tail of the distribution are expanding more than proportionally than the middle-ranked products, reducing the deviation from the straight line in log scale predicted by the Pareto distribution in equation (3). Figure 4 shows the evolution of these coefficients by country. Although with different trajectories, all 17 countries in the sample show a decreasing curvature. Again, following Bar-Isaac et al. (2012), this is consistent with a scenario in which search costs are reduced, information flows more easily among consumers and both superstars and the long-tail increase their shares of sales.

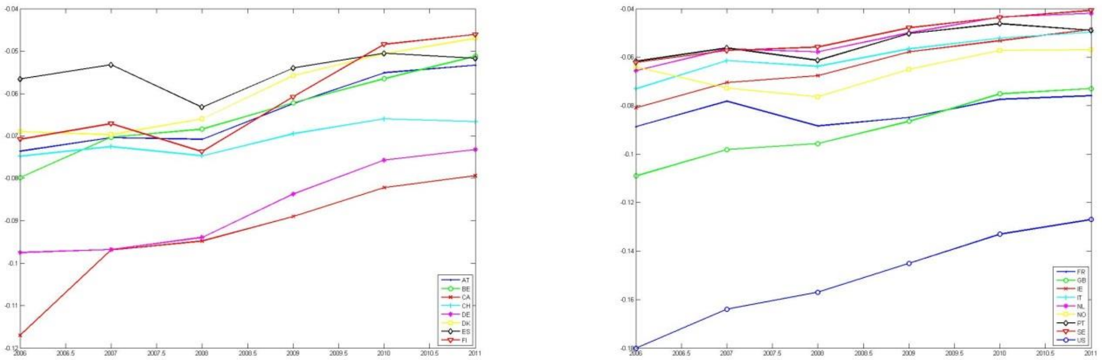
Table 5: Estimates of the Rank-Downloads relationship by country

	Austria	**Belgium**	**Canada**	**Switzerland**	**Germany**	**Denmark**	**Spain**	**Finland**	**France**
ln(rank)	-0.0342***	-0.103***	0.292***	0.149***	0.376***	-0.209***	-0.243***	-0.129***	0.244***
	(0.00308)	(0.00111)	(0.00190)	(0.00289)	(0.00359)	(0.00140)	(0.00449)	(0.00230)	(0.00253)
$[\ln(\text{rank})]^2$	-0.0581***	-0.0566***	-0.0772***	-0.0646***	-0.0769***	-0.0550***	-0.0491***	-0.0534***	-0.0727***
	(0.000143)	(4.97e-05)	(7.78e-05)	(0.000126)	(0.000148)	(6.46e-05)	(0.000205)	(0.000115)	(0.000106)
Constant	7.850***	9.238***	8.879***	7.679***	7.793***	9.589***	9.100***	7.283***	8.016***
	(0.0165)	(0.00622)	(0.0115)	(0.0165)	(0.0217)	(0.00760)	(0.0244)	(0.0115)	(0.0150)
Observations	2,392,627	3,272,883	8,071,002	4,484,358	8,096,139	2,477,841	2,722,095	1,183,042	6,267,545
R-squared	0.999	0.998	0.996	0.998	0.996	0.999	0.997	0.998	0.997

	United Kingdom	**Ireland**	**Italy**	**Netherlands**	**Norway**	**Portugal**	**Sweden**	**United States**
ln(rank)	0.248***	-0.201***	-0.241***	-0.259***	0.0174***	-0.140***	-0.264***	1.557***
	(0.00211)	(0.00108)	(0.000739)	(0.00112)	(0.00338)	(0.00450)	(0.00161)	(0.00389)
$[\ln(\text{rank})]^2$	-0.0737***	-0.0567***	-0.0505***	-0.0456***	-0.0626***	-0.0531***	-0.0459***	-0.128***
	(8.39e-05)	(5.04e-05)	(3.23e-05)	(5.06e-05)	(0.000155)	(0.000232)	(7.45e-05)	(0.000144)
Constant	9.728***	9.276***	9.859***	9.383***	7.837***	6.770***	9.164***	4.752***
	(0.0132)	(0.00575)	(0.00421)	(0.00619)	(0.0183)	(0.0217)	(0.00871)	(0.0262)
Observations	11,616,435	1,955,614	3,708,194	3,059,281	2,590,004	892,679	2,582,481	27,425,471
R-squared	0.994	0.999	0.999	0.999	0.999	0.999	0.999	0.988

Note: all specifications include time dummies. Robust standard errors in parentheses. *** denote significant at 1%, ** at 5% and * at 10%, respectively

Figure 4: Evolution of rank-downloads curvature coefficient 2006-2011



a) Countries (sample)

b) Countries (sample)

In addition, Johnson and Myatt (2006) have shown that demand transformations result from changes in the dispersion of consumer's valuations which lead to rotations of the demand curve. One driver of these changes can be associated with information provision. As the authors argue, there are two potential functions of information provision related activities. First, they can be of a promotional type, highlighting the product's existence and shifting the demand curve outwards. Second, they can provide real information to consumers, making the demand curve to rotate.

In order to explain the decay in the curvature of the rank-downloads relationship we model the effects of increasing information transmission in electronic markets and the availability of different search and recommendation technologies on downloads. Following the literature (Brynjolfsson et al., 2006; Chevalier and Golsbee, 2003; Ghose and Gu, 2006, Hinz et al., 2011), we extend our model represented in equation (4) that relates the log of sales to the log of demand ranks and its square term. In equation (4), α represents the height of the parabola, or the point of intersection with the y-axis. Hence, changes in this parameter are related to shifts in the downloads-rank relationship. Moreover, β measures the declivity of the parabola and γ controls the speed of the decrease of the function from the vertex. Hence, changes in both β and γ represent rotations of the downloads-rank relationship. We analyse the effects of search costs and information transmission on the distribution of demand by linking the parameters α, β and γ with the drivers described in what follows.

In order to capture the effects of search costs and information transmission on sales concentration in digital music we average by country and year the popularity measure given by Google trends for four different information sources. The first is *iTunes*, the main online music retailer, as a proxy for search costs. The next one is *Facebook*, the most popular social network, is used here to control for online word-of-mouth (eWOM) through social interactions between consumers and peers. Our third source of information is *LastFM*, one of the most important music webpages and known for its algorithm designed to help users to find their preferred matches, and its inclusion is justified by the idea of controlling for the usage of recommendation systems. Lastly, we also include *Youtube* as a source of information, representing a quasi-centralised promotion platform since many media corporations use it to offer some of their music related material.

In addition to dummy variables reflecting the intensity of the popularity of each information source described and that take into account the shifts in the demand relationship (alpha), we include their interactions with ln(rank) and $\ln(\text{rank})^2$ to capture possible changes in the slope of the download-rank relationship that would represent rotations of the demand relationship. To avoid excessive multi-collinearity, ln(rank) is centred before computing its squared term or the interactions. With this procedure, the post-estimation collinearity tests are reasonable and the estimations consistent among different specifications and robustness checks. In addition, heteroskedasticity is tackled by computing robust standard errors. Columns (1) to (4) in Table 6 show the results when introducing the information sources sequentially in equation (4). The last column of the table presents the results when all sources of information are included in equation (4) at the same time.

The first column of Table 6 shows the results related to iTunes. From the table is evident that the more popular iTunes is, it shifts the downloads-rank relationship outwards and, at the same time, increases the speed at which the curve changes. Hence, it has clearly a superstar effect, inducing more sales concentration. The second column offers the results for Facebook. Here, eWOM also shifts the demand relation outward, and although the slope of the parabola is smaller, the speed at which the function decreases is higher. In this case, eWOM would tend to favour superstar and middle-ranked songs and to penalise the long-tail. As stated before, Lastfm represents the popularity of recommender systems. Recently, Hervas-Drane (2013) suggested that sales concentration should be reduced in the presence of recommendation systems given that consumers will be able to find better matches to their tastes. Table 6 shows that this is indeed the case: Lastfm not only shifts the demand relation inward but decrease both the slope and the speed at which the function decreases. This means an important counter clockwise rotation, favouring the long-tail. Finally, the results from Youtube favour both the superstars and the long tail since it makes the demand relation to shift outward and at the same time reduces the slope and the speed of decrease.

Table 6: Determinants of shifts and rotations in the rank-downloads relationship

	(1)	(2)	(3)	(4)	(5)
$\ln(rank_c)$	-1.685***	-1.759***	-1.756***	-1.800***	-1.799***
	(0.000113)	(6.54e-05)	(6.19e-05)	(5.00e-05)	(0.000169)
$[\ln(rank_c)]^2$	-0.0760***	-0.0926***	-0.0971***	-0.103***	-0.0980***
	(4.02e-05)	(4.17e-05)	(5.13e-05)	(4.71e-05)	(7.81e-05)
iTunes	0.0505***				-0.0225***
	(0.000123)				(0.000154)
$\ln(rank_c)$*iTunes	-0.0509***				-0.0476***
	(0.000114)				(0.000143)
$[\ln(rank_c)]^2$*iTunes	-0.0416***				-0.0271***
	(9.64e-05)				(8.86e-05)
Facebook		0.0497***			0.0388***
		(0.000156)			(0.000133)
$\ln(rank_c)$*Facebook		0.0329***			0.0407***
		(4.82e-05)			(4.65e-05)
$[\ln(rank_c)]^2$*Facebook		-0.0121***			0.00143***
		(7.06e-05)			(5.14e-05)
Lastfm			-0.0555***		-0.0564***
			(0.000144)		(0.000127)
$\ln(rank_c)$*Lastfm			0.0323***		0.0613***
			(4.80e-05)		(4.37e-05)
$[\ln(rank_c)]^2$*Lastfm			0.00171***		0.00766***
			(6.45e-05)		(5.36e-05)
Youtube				0.107***	0.0947***
				(0.000156)	(0.000155)
$\ln(rank_c)$*Youtube				0.196***	0.165***
				(8.17e-05)	(0.000148)
$[\ln(rank_c)]^2$*Youtube				0.0356***	0.0369***
				(6.04e-05)	(7.16e-05)
Constant	-2.725***	-2.759***	-2.699***	-2.750***	-2.719***
	(0.000147)	(0.000150)	(0.000157)	(0.000173)	(0.000161)
Observations	92,797,691	92,797,691	92,797,691	92,797,691	92,797,691
R-squared	0.985	0.984	0.984	0.986	0.987

Note: estimations include country and time fixed effects. *** denote significant at 1%, ** at 5% and * at 10%, respectively.

Our results in Table 6 show that changes in search costs and information transmission technologies significantly affect the shape of the distribution by means of both shifts and, more importantly, rotations. In some cases, these changes positively impact the superstars and hence provoke more sales concentration. In other cases, the long-tail is promoted and thus sales concentration is reduced. Except for recommender systems (LastFM), all the information sources used promote more sales concentration by means of outward shifts of the demand relation. In contrast, both LastFM and Youtube promote unambiguously as well the long-tail, by rotating the demand relation through a decreasing slope and less concavity. In addition, the results when we consider all four information sources together (column 5) are –with minor changes with respect to the previous columns- similar to those already discussed.

When analysing how information provision/exchange affect demand, Johnson and Myatt (2006) identify two different functions. On one hand, promotional hype, which highlights the product existence causing the demand curve to shift. On the other hand, the provision of real information, which rotates the demand curve. In general, more information allow consumers to learn how good is the match between their preferences and the product attributes, and hence their valuation for it. Under this setting, this need not always increase demand, as some consumers will learn that the product is not suited to their tastes. However, if more information does not reduce search costs significantly, or if variety increases abundantly, consumers might simply follow recommendations to avoid lengthy and costly search processes.

In summary, we observe different effects. With some online information tools we observe the emergence of a typical long-tail demand distribution pattern such that demand becomes more evenly distributed and superstars lose importance. With others, we observe more sales concentration favouring the superstars. These differences in information transmission mechanisms act trough different channels. The reduction of search costs do not necessarily minimise the importance of superstars. We observe that with the use and proliferation of some of these online tools consumers' preferences move towards the extremes and middle-ranked products suffer lost demand. This is consistent with the theoretical contributions of Johnson and Myatt (2006), Bar-Isaac et al. (2012) and Hervas-Drane (2013) where information flows that reduce search costs tend to generate both shifts and rotations in the demand curve and hence promoting simultaneously the occurrence of both superstar and long-tail effects.

7 Conclusions

It is often assumed that the internet brings benefits to consumers through several channels, including lower prices and more variety of consumer products to choose from – known as the long tail effect (Anderson, 2006). Considerable empirical evidence has built up in the past decade in support of the variety effect in specific markets. However, some authors have emphasized that search costs may block the realization of this long tail effect. Confronted with information overload, consumers use the evaluations of others as an indicator of product quality.

In this paper we examine the variety hypothesis in the entire online market for digital music downloads, including 17 countries over the period 2006-2011. A wide range of online sales platforms and discovery channels play a role in this market. First, we have characterised the distribution of legal music downloads and find that it is heavily skewed despite the predicted long-tail effect due to the expansion of online information tools that help consumers to find products that best match their preferences. Second, we test the effect of the presence of these online information platforms as an indication of reduced search costs on demand. We show that some of these tools generate a superstar effect, while others tend to produce a long-tail pattern.

We hypothesize that this process of online information transmission is due to a wider availability and variety of search tools. Ubiquitous and very popular word-of-mouth search and discovery channels such as Facebook and sales reinforcing discovery channels such as iTunes sales rankings tend to push consumers towards the superstar songs and artists, even in mature markets. On the other hand, specialised music recommendation systems - like Lastfm- make the demand curve to rotate counter clockwise, favouring the long-tail and penalising superstars. The net effect is ambiguous since many information tools and discovery channels are present in the electronic music market. In general terms, the main conclusion of this paper resides in the co-existence of both the superstar and the long-tail effects. We have shown that online information transmission tools are a source of demand transformations, some of them shifting demand and some of the demand curve to rotate. Further research will be required to find out what type of search cost reducing discovery channels can explain the growing long tail effect.

This research focuses on the market for digital music downloads only. We observe the evolution of the distribution of sales in that market over a relatively short time period since it emergence in 2006 and up to a more mature stage in 2011. It would of course be interesting to compare this with the distribution of sales in pre-online physical music sales markets. Unfortunately, there are no data available on the long tail distribution in physical music sales. The best data sets do not go further than the top-500 or in exceptional cases the top-1000 album sales. For digital sales we are for the first time able to observe the real long tail at individual song level composed of millions of songs.

Note that we do not attempt to estimate consumer welfare effects. We have not tried to estimate the marginal contribution to consumer welfare of superstar and long tail sales. Consequently we do not know at this stage to what extent changes in the distribution of sales affect consumer welfare. Consumer welfare is however a primary concern for policy makers. To the extent that changes in the sales distribution are affected by search costs and information discovery channels, policy makers should keep an eye on the market for these discovery channels and ensure sufficient competition in these markets. In our data set we find that the larger and more ubiquitous discovery channels do not necessarily favour the long tail. Concentration in the discovery channel market would reinforce superstars at the expense of the long tail.

References

Adler M. (1985), Stardom and Talent, American Economic Review, 75(1): 208-212.

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

Alexander P. J. (1996), Entropy and Popular Culture: Product Diversity in the Popular Music Recording Industry, American Sociological Review, 61(1): 171-174.

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

Bar-Isaac, H., G. Caruana and V. Cuñat (2012), Search, Design, and Market Structure. American Economic Review, April 2012, 102(2): 1140–1160.

Baye, M., B. de los Santos and M. R. Wildenbeest (2013a) The Evolution of Product Search. Journal of Law, Economics and Policy, 9: 201-221.

Baye, M., B. de los Santos and M. R. Wildenbeest (2013b), Searching for Physical and Digital Media: The Evolution of Platforms for Finding Books. In: NBER Economics of Digitization, edited by Shane Greenstein, Avi Goldfarb, and Catherine Tucker. Chicago: University of Chicago Press.

Bakos, Y. (1997), Reducing buyer search costs: implications for electronic marketplaces. Management Science, 43(12): 1676–1692.

Brynjolfsson, E., Y. Hu and M. Smith (2003), Consumer surplus in the digital economy: estimating the values of increased product variety at online booksellers. Management Science, 49(11): 1580–1596.

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

Brynjolfsson, E., Hu, Y., and Smith, M., (2010), Long tails vs. superstars: the effect of information technology on product variety and sales concentration patterns. Information Systems Research, 21(4): 736–747.

Brynjolfsson, E., Hu, Y., and Simester, D. (2011), Goodbye Pareto principle, hello long tail: the effect of search costs on the concentration of product sales. Management Science, 57(8): 1373–1386.

Bourreau, M., F. Moreau, M. Gensollen and P. Waelbroeck (2013), Selling less of more? The impact of digitization on record companies. Journal of Cultural Economics, 37(3): 327-346.

Burnett R. (1992), The Implications of Ownership Changes on Concentration and Diversity in the Phonogram Industry, Communication Research, 19: 749-769.

Crain, W. and R. Tollison (2002), Consumer Choice and the Popular Music Industry: A Test of the Superstar Theory," Empirica, 29(1): 1-9.

Chen, Y, Y. Wang and J. Xie (2011), Online Social Interactions: A Natural Experiment on Word of Mouth versus Observational Learning. Journal of Marketing Research, XLVIII: 238 –254.

Chevalier, J. and A. Goslbee (2003), Measuring Prices and Price Competition Online: Amazon.Com and Barnesandnoble.Com. Quantitative Marketing and Economics, 1(2): 203-222.

Chevalier J. and D. Mayzlin (2006), The Effect of Word of Mouth on Sales: Online Book Reviews, Journal of Marketing Research, 43(3): 345-354.

Chung K.H. and R.A.K. Cox (1994), A Stochastic Model of Superstardom: An Application of the Yule Distribution, Review of Economics and Statistics, 76: 771-775.

Clauset, A., C. R. Shalizi and M. E. J. Newman (2009), Power-law distributions in empirical data. SIAM Review 51(4): 661 – 703.

Clemons, E., Gao, G., and Hitt, L. (2006), When online reviews meet hyperdifferentiation: a study of the craft beer industry. Journal of Management Information Systems, 23(2): 149–171.

Dewan, S. and J. Ramaprasad (2012), Music Blogging, Online Sampling, and the Long Tail. Information systems Research, 23(3) part-2: 1056-1067.

De Los Santos, B., A. Hortaçsu and M. R. Wildenbeest (2012), Testing Models of Consumer Search Using Data on Web Browsing and Purchasing Behavior. American Economic Review, 102(6): 2955-2980

Dowd T. J. (2004), Concentration and Diversity Revisited: Production Logics and the U.S. Mainstream Recording Market, 1940 to 1990, Social Forces, 82: 1411-1455.

Duan, W., Gu, B., and Whinston, A. B. (2008a), The dynamics of online word-of-mouth and product sales—an empirical investigation of the movie industry. Journal of Retailing, 84(2): 233–242

Duan, W., B. Gu and A. B. Whinston (2008b) Do online reviews matter? — An empirical investigation of panel data. Decision Support Systems 45: 1007–1016.

Duan, W., Gu, B., and Whinston, A. B. (2009), Informational cascades and software adoption on the internet: an empirical investigation. MIS Quarterly, 33(1): 23–48.

Elberse, A., and Oberholzer-Gee, F. (2007) Superstars and underdogs: an examination of the long tail phenomenon in video sales. Working paper no.07-120, Harvard Business School.

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

Elliott, C. and R. Simmons (2011), Factors determining UK Album Success. Applied Economics, 43(30): 4699-4705.

Giles, D. (2007), Increasing returns to information in the US popular music industry. Applied economic letters 14(5-6): 327-331.

Goldmanis, M., Hortaçsu, A., Syverson, C., and Emre, O. (2009), E-commerce and the market structure of retail industries. The Economic Journal, 120(545): 651–682.

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

Ghose, A. and B. Gu (2007), Estimating Menu Costs in Electronic Markets. In: Proceedings of the International Conference on Information Systems, ICIS 2007, Montreal, Quebec, Canada.

Hamlen W. (1991), Superstardom in Popular Music: Empirical Evidence, Review of Economics and Statistics, 73(4): 729-733.

Hamlen W.A. (1994), Variety and Superstardom in Popular Music, Economic Inquiry, 32: 395-406.

Hervas-Drane, A. (2013), Word of Mouth and Sales Concentration, NET Institute Working Paper No. 07-41.

Hand, C. (2001), Increasing Returns to Information: Further Evidence from the UK Film Market, Applied Economics Letters, 8: 419-421.

Hinz, O., J. Eckert and B. Skiera (2011), Drivers of the long tail phenomenon: an empirical analysis. Journal of Management Information Systems, 27(4): 43-69.

IFPI (2013) Recording Industry in numbers. The recorded music market in 2012. London.

Johnson, J. P. and D. P. Myatt (2006), On the Simple Economics of Advertising, Marketing, and Product Design. American Economic Review, 96(3): 756-784.

Ijiri, Y. and Simon, H. (1974), Interpretations of Departures from the Pareto Curve Firm-Size Distributions, Journal of Political Economy, 82(2): 315-332.

Kendall, T. D. and K. Tsui (2011), The Economics of the Long Tail. The B.E. Journal of Economic Analysis & Policy Topics, Vol. 11(1): article 76.

Krueger A.B. (2005), The Economics of Real Superstars: The Market for Concerts in the Material World, Journal of Labor Economics, 23(1): 1-30.

Kumar, A., M. D. Smith and R. Telang (2011), Information Discovery and the Long Tail of Motion Picture Content, mimeo.

Liu, Y. (2006) Word of mouth for movies: its dynamics and impact on box office revenue. Journal of Marketing, 70(3): 74–89.

Lopes P.D. (1992), Innovation and Diversity in the Popular Music Industry, 1969 to 1990, American Sociological Review, 57(1): 56-71.

McKenzie, J. (2008), Bayesian Information transmission and stable distributions: motion picture revenues and the Australian box office. Economic Record 84(266): 338-353.

Newman, M. E. J. (2005), Power laws, Pareto distributions and Zipf's law. Contemporary Physics 46: 323–351.

Oestreicher-Singer, G. and A. Sundararajan (2012), Recommendation networks and the long tail of electronic commerce. MIS Quarterly 36(1): 65-83.

Page, W. and E. Garland (2009), The long tail of P2P. PRS for Music Economic Insight, Issue 14.

Pareto, V. (1897) Cours d'economie politique, Rouge, Paris.

Peterson R.A. and D.G. Berger (1975), Cycles in Symbol Production: the Case of Popular Music, American Sociological Review, 40(2): 158-173.

Peterson R.A. and D.G. Berger (1996), Measuring Industry Concentration, Diversity, and Innovation in Popular Music, American Sociological Review, 61(1): 175-178.

Rosen S. (1981), The Economics of Superstars, American Economic Review, 71(5): 845-58.

Simon, H. (1955), On a class of skew distribution functions. Biometrica, 52:425-440.

Strobl E. and C. Tucker (2000), The Dynamics of Chart Success in the UK Pre-recorded Popular Music Industry, Journal of Cultural Economics, 24(2): 113-134.

Tan and Netessine 2009 Is Tom Cruise Threatened? Using Netfix Prize Data to Examine the Long Tail of Electronic Commerce. Mimeo.

Tucker, C. and J. Zhang (2007), How Does Popularity Information Affect Choices? A Field Experiment. Management Science, 57(5): 828–842.

Weeds, H. (2012), Superstars and the Long Tail: The Impact of Technology on Market Structure in Media Industries. Information Economics and Policy, 24(1): 60-68.

Zhao, X., Gu, B., and Whinston, A. B. (2008), The influence of online word-of-mouth long tail formation: an empirical analysis. In Proceedings of the Conference on Information Systems and Technology (CIST 2008), Washington, DC, October 11–12.

Zhou, W. and W. Duan (2012), Online user reviews, product variety, and the long tail: An empirical investigation on online software downloads. Electronic Commerce Research and Applications 11: 275–289.

Zhu, F. and M. Zhang (2010), Impact of Online Consumer Reviews on Sales: The Moderating Role of Product and Consumer Characteristics. Journal of Marketing, 74: 133–148.

European Commission
Joint Research Centre – Institute for Prospective Technological Studies

Title: **Search costs, information exchange and sales concentration in the digital music industry**

Author(s): Néstor Duch-Brown and Bertin Martens

Spain: European Commission, Joint Research Centre

2014 – 25 pp. – 21.0 x 29.7 cm

Scientific and Technical Research series – ISSN 1831-9408 (online)

Abstract

It is often assumed that consumers benefit from the internet because it offers a “long tail” with more variety of products to choose from. However, search costs may block the long tail effect and result in the dominance of superstars. This paper examines the variety hypothesis in the entire online market for digital music downloads in 17 countries over the period 2006-2011. First, we show that the entire distribution of legal music downloads is heavily skewed. Second, we hypothesise that a wide range of online information channels (sales and discovery platforms) play a role in this market. We find that the reduction of search costs implied by the generalisation of online information tools transforms demand as a result of changes in the dispersion of preferences. Ubiquitous and very popular discovery channels such as Facebook and iTunes tend to push consumers towards the superstars by shifting the demand curve but also towards the long-tail since they also generate rotations that promote niches. Consequently, both the superstar and the long tail effects emerge even in mature digital markets.

JRC Mission

As the Commission's in-house science service, the Joint Research Centre's mission is to provide EU policies with independent, evidence-based scientific and technical support throughout the whole policy cycle.

Working in close cooperation with policy Directorates-General, the JRC addresses key societal challenges while stimulating innovation through developing new methods, tools and standards, and sharing its know-how with the Member States, the scientific community and international partners.

Serving society
Stimulating innovation
Supporting legislation