WHERE DO ARTISTS PERFORM CONCERTS IN AGE OF FILE SHARING?

Completed Research Paper

Daegon Cho
Heinz College, Carnegie Mellon University
4800 Forbes Ave., Pittsburgh, PA, USA 15213
daegonc@andrew.cmu.edu

Abstract

While recorded music sales have plummeted due to the prevalence of file sharing, live performances have become an important revenue source for artists. The concert industry has grown steadily, as demonstrated by the increase in the number of concerts and performing artists. This paper studies how concert distribution has evolved between well-known artists (Superstars) and lesser-known artists (Underdogs) in the era of file sharing. We first develop an artist's time allocation model to explain the underlying phenomenon, suggesting that artists would perform more concerts and the distributional trend might be different across popularity. To test this model, we empirically examine the distribution patterns over time. Our findings indicate that concert distribution of Superstars follows the Long Tail effect and that the strengthened Pareto curve is observed from that of Underdogs. Demand- and supply-side drivers related to the impact of the Internet would lead to this distinctive difference.

Keywords: Concert, Superstars, Long Tail, Power Law Distribution

Introduction

The traditional structure and business models in the music industry have been altered due to advancements in information technology, and two linked phenomena lie at the heart of these changes over the past decade. First, in recent years, there have been more opportunities for artists to take control of their own careers, whereas there were few opportunities for an artist to become successful without signing a contract with a music label in the past (e.g., YouTube has replaced MTV, and social networking sites allow a new level of communications and interactions with fans.) The viral potential of social media creates the possibility to become well known and to promote a new album release and live tours with lower costs. The second development in the music industry has been an important, although for some, a difficult, change of income source for artists. Music artists typically generate income in two ways: royalties from recorded music sales and revenue sharing from live performances. They used to earn millions of dollars by selling records, because buying a physical copy of an album was a primary way to listen to music. However, the digital format (e.g., MP3) of a song has replaced the hard copy of an album, and file sharing has been flourishing since the advent of Napster in 1999.

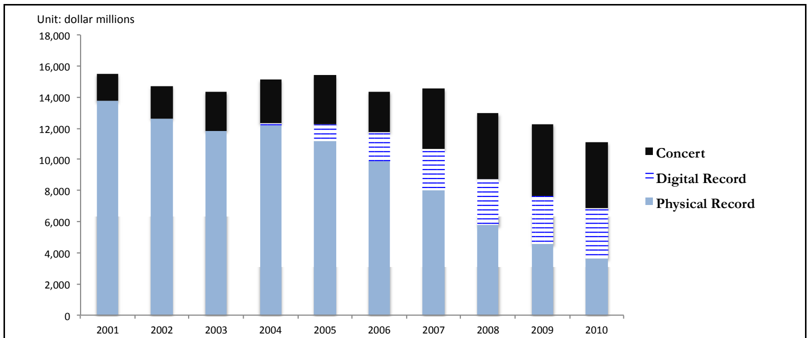

Figure 1. Sales Trend of the Music Industry in the U.S. (2001-2010)[1]

Figure 1 depicts the sales trend of the U.S. music industry in the period from 2001 to 2010, indicating that a sharp drop in recording revenues comes hand in hand with the rise of the live performance market as a growing source of income. The market size of the concert market has risen from $1.75 billion in 2001 to $4.25 billion in 2010, as opposed to the striking decline in the sales of physical recordings in the same period. As Connolly and Krueger (2006) and other industry experts suggested, this shift would provide a potentially important opportunity for artists, live performances, but there has been little academic research to focus on the dynamics of the growing concert market. In other words, previous research on the music industry has been limited to only a fraction of the industry – the recording segment – while ignoring a quickly rising segment, that is, live concerts. Our paper studies this important trend by focusing on concert location distribution in the post-file sharing age. We compare the distribution and concentration of concert locations over the past decade for two groups of artists: those who are well-known Superstars and those who are lesser-known Underdogs.[2] Our work relates to three strands of literature:

The first focuses on the impact of piracy, suggesting the declining sales of physical information goods are related to the file sharing. Economists have studied the theoretical aspects of copying and the network effects resulting from digital piracy (Dejean, 2009). To explore the relationship between file sharing and revenues in the industry, earlier empirical studies estimated the impact and the consequence of digital piracy in different contexts. By using country level data, Peitz and Waelbroeck (2004) and Zentner (2005) estimated the reduction in music sales caused by broadband penetration, whereas Boorstin (2004) reported a positive relationship between Internet access and CD sales, so-called sampling effects. More recent empirical work was accomplished by using more detailed data of P2P activities (Oberholzer-Gee and Strumpf, 2007; Blackburn 2007; Smith and Telang, 2009); their results showed little effect of file sharing on the sales of physical information goods. However, other studies using survey-based data generally found a negative impact of P2P (Michel 2006; Hong 2005).

[1] Source: Recording Industry Association of America (RIAA), Pollstar Magazine

[2] More detailed descriptions of these two groups will be presented in the following section.

A second important literature examines the link between the Superstar effect and advances in technology. As Rosen (1981) famously demonstrated, the small number of artists represents a large portion of sales by predicting that lower transaction costs will homogenize patterns of consumption. This argument is consistent with technological advances because technology allows the most talented performers to be exposed to more consumers, and as a result, a few superstars are more likely to dominate the market. A large body of empirical studies has conducted based on this theory.

A third, and final, strand of literature upon which we focus is the Long Tail effect led by information technology. This term was coined by Anderson (2006) to describe how aggregated sales of niche products in online commerce can contribute a large share of sales. However, the value of Long Tail has been in dispute in subsequent academic research. Some researchers found evidence that the Internet would lead to higher concentration by popular products on the information goods market (Elberse and Oberholzer-Gee, 2008; Tan and Netessine, 2011), whereas others documented the existence of the Long Tail phenomenon by identifying different drivers from the supply side (lower inventory and distribution costs) and from the demand side (easy search tools and useful recommender systems to access niche products) (Brynjolfsson et al., 2006; Brynjolfsson et al., 2010; Brynjolfsson et al., 2011; Fleder and Hosanagar, 2009). The concert industry also features these kinds of factors and we focus on the narrow, but important, question: how has concert location distribution evolved in the digital age?

As noted above, to the best of our knowledge there has been little academic work on this topic, and a key reason for this limitation is the lack of data.[3] A distinctive feature of our paper is the use of comprehensive and up-to-date concert data, which accurately records the date and place of performances by each artist. This data is matched with the Core Based Statistical Area (CBSA), the standard for the U.S. geographic area, which is defined by the Office of Management and Budget. The division is defined by an urban center of at least 10,000 people and adjacent areas that are socioeconomically tied by commuting. The resulting data covers all of the concerts performed by 45 Superstars and 224 Underdogs in the U.S. in the period from 2000 to 2011; our dataset includes more than 70,000 observations of concerts, which allows us to investigate the distributional dynamics in terms of concert location as a result of the artist's choice. Our empirical findings suggest that concert distribution of Superstars follows the Long Tail effect and that the strengthened Pareto curve is observed from that of Underdogs. This distinctive feature implies that demand- and supply-side drivers related to the impact of the Internet would lead to this distinctive difference.

The remainder of our paper proceeds as follows: In Section 2, we present a model in which artists make a decision to perform live tours and choose a location to maximize their utility. The model provides the logic that drives artists to perform more concerts and to choose a location for live tours in the given environment in which their income from recording sales has declined. That is, our model shows that, to maximize their utility, it is necessary to increase their live performances. Sections 3 and 4 report empirical exercise. In Section 3 we describe our data and in Section 4 we present our findings about the distribution dynamics across popularity in the 2000s. Section 5 provides a conclusion.

A Model of the Artist's Time Allocation

To fix our ideas and set the stage for the empirical analysis to follow, in this section we provide a time allocation model for artists based on the seminal work by Becker (1965). At the core of our model is the assumption that an artist's utility, *U*, depends on income, *y*, and leisure, *L*. We thus write $U = f(y, L)$, assuming that both income and leisure can be regarded as "goods" and $f(y, L)$ is strictly concave and increasing in both income and leisure, with $f_y(y, L) > 0$ and $f_L(y, L) > 0$. We also expect $f_{yy}(y, L) < 0$ and $f_{LL}(y, L) < 0$. Suppose that an artist generates her income, *y*, by selling records and performing concerts. Let T_R and T_C be the amount of time for recording and concerts, respectively, and she devotes the remainder of her time for leisure and other activities. The artist can tour in different locations, a_i $(i = 1, ..., N)$, and allocate T_C appropriately. Each location can be ordered according to its attractiveness. All else being equal, large regions with a higher population density (e.g., New York City, NY) are preferable to smaller regions (e.g., Albany, NY). Not only is the demand higher, but also the infrastructure for

[3] One exception is a study by Mortimer et al. (2012) in which they empirically examined the relationship between recorded music sales and concert revenues by using concert data from 1993 to 2002. However, the period the data covered is too early to measure all of the dynamics of the concert market that have been led by the Internet, and the study did not provide details of trends of concert location distribution.

performing concerts (e.g., affordable venues) is well established in such regions. For example, those areas have a variety of venues ranging from small concert halls to large stadiums, which can accommodate the expected audience. Accordingly, a well-known artist will begin from the most attractive location and go down the list of $a_i \ (i = 1,..., N^/)$, and stop at $N^/$ when she reaches her optimal T_C. In this context, we are interested in comparative statics of what happens when recording incomes are declining due to changing technology and P2P file sharing.

We further assume that "recording" is a must for concerts to remain viable. In other words, independent of how much income an artist can make from recording sales, she has to invest that time to generate incomes from concerts. Thus, T_R remains unchanged even if file sharing has reduced recording revenues substantially. Formally suppose that the artist receives income, *R*, from recording and receives $h \times T_C$ from concerts, where *h* is income per unit time from a concert. Thus, the income is

$$y = hT_C + R, \tag{1}$$

and the constraint on time is

$$T = T_R + T_C + L \tag{2}$$

where *T* is the total available time for the artist. Since T_R is not a choice variable, we can simply normalize Equation (2) into $T = T_C + L$, or $L = T - T_C$. Thus, the artist chooses T_C to maximize her utility as follows:

$$Max \ U = f(y, L) = f(hT_C + R, \ T - T_C) \tag{3}$$

Taking the first order condition with respect to T_C from Equation (3) yields

$$f_y(y, L) + f_L(y, L) = h \cdot f_y(hT_C + R, T - T_C) - f_L(hT_C + R, T - T_C) = 0 \tag{4}$$

where f_y and f_L are derivative. For an interior optimal maxima, the second order condition must be satisfied. Thus,

$$h^2 f_{yy}(\cdot) - hf_{yL}(\cdot) - hf_{Ly}(\cdot) + f_{LL}(\cdot) < 0. \tag{5}$$

After file sharing is widely available, the income from recording sales, *R*, has decreased. How does this change affect the artist's decision to invest in concerts? In short, we will explore the sign of dT_C / dR. To do this, we take a full derivative of Equation (4), and it yields:

$$\begin{gathered} \left[h^2 f_{yy}(\cdot) - hf_{yL}(\cdot) - hf_{Ly}(\cdot) + f_{LL}(\cdot)\right] dT_C + \left[hf_{yy}(\cdot) - f_{Ly}(\cdot)\right] dR = 0 \\ or \\ \frac{dT_C}{dR} = -\frac{hf_{yy}(\cdot) - f_{Ly}(\cdot)}{h^2 f_{yy}(\cdot) - hf_{yL}(\cdot) - hf_{Ly}(\cdot) + f_{LL}(\cdot)} \end{gathered} \tag{6}$$

The denominator in Equation (6) is negative from Equation (5), and the first term in the numerator, $hf_{yy}(\cdot)$, is negative by assumption. Thus, the sign of dT_C / dR depends on the sign of the second term of the numerator, $f_{Ly}(\cdot)$. If $f_{Ly}(\cdot)$ is positive (or not too negative), then the numerator is negative. $f_{Ly}(\cdot)$ is a marginal utility of leisure when the income increases. As long as the marginal utility of leisure increases (or does not decrease too much) according to the increase in income, $f_{Ly}(\cdot)$ is positive. In fact, if leisure is a "normal" good, then this is always true. Therefore, when *R* goes down under reasonable assumptions on artist's utility, we expect that T_C will rise. In short, artists will have an incentive to spend more time on concerts given that file sharing decreases recording sales. This extra time will be spent on touring locations $a_i \ (i = N^/,..., N)$ which the artist did not tour before.[4]

This leads to our first hypothesis:

[4] We assume that touring the same location repeatedly within the same period is not a profitable strategy.

> *HYPOTHESIS 1: Artists are more likely to perform more concerts after file sharing.*

Let us take the popularity into account in this context. Since a Superstar will visit smaller locations due to the necessity of increasing her efforts in concerts, we expect the distribution of concerts along with geographic locations to flatten out. For example, Superstars tended to tour only larger cities in the past; therefore, those larger cities may constitute a large share of the total number of concerts in the given period. As Superstars tour more in smaller places, the share of large cities should decline. Notice that we are not arguing that the absolute number of concerts in larger cities will decline. Instead, we are simply stating that the share of concerts from the larger cities will decline.

> *HYPOTHESIS 2: For Superstars, geographic distribution of concerts will become more dispersed (or follow the Long Tail trend in terms of the location) after file sharing.*

Thus far, the story for Superstars is mostly from the supply side. That is, they are already well known, so the sampling effect from file sharing is limited. For example, file sharing may increase Superstars' popularity in smaller towns, but mostly the effect is likely to be marginal. They have an incentive to tour smaller places to recoup the money they have lost in recording sales.

However, for lesser-known artists (Underdogs), we note that their story is mostly from the demand side. Suppose that a lesser-known artist initially originates from her hometown, so she may not encounter a demand for live performance outside her local region. Therefore, even though there would be a potential demand for her concerts in big cities, she tends to stay in her local region due to the uncertainty of success in the big city. In other words, the distribution of concerts for Underdogs would be more scattered than that for Superstars.[5] However, the impact from the Internet, e.g., free music distribution, file sharing, video clips on YouTube, and more frequent interactions with fans on Facebook and Twitter, should provide greater opportunities for Underdogs to be better known in a greater number of regions and to reduce the risk of failures in touring. Thus, they may be able to tour new places where they had not toured previously. If this is indeed the case, more Underdogs would have newly-created demand in larger cities and hence the distribution of concerts should shift more toward larger cities than before.

> *HYPOTHESIS 3: For Underdogs, geographic distribution of concerts will become more concentrated (or follow the Pareto trend in terms of the location) after file sharing.*

In summary, we expect the distribution of concert curves to shift in opposite direction depending on whether an artist was well known or not.

Data

Source and Sampling Process

Our paper hinges on access to a unique data source. Our dataset is historical concert information including performers, date, place (location and venue), which came from Songkick[6], a website that provides information of live music events. We collected concert information of the period from 2000 to 2011 at the artist-level.

In order to compare the trends of concert location between Superstars and Underdogs, it is important to divide the two groups of artists appropriately. Because we focus our analyses on the period for 12 years, there might be a possibility that a band that was not well known in the early period became more popular later in our study period or vice versa. By taking this difficulty and a possible bias into account, we construct two subsets of data as following procedures. First, there are two criteria to classify the group of "Superstars" by using the top 100 tours in North America provided by Pollstar Magazine, which is one of the most trustable trade magazines in the concert industry. On the one hand, Superstars are defined as artists who generate a sufficient large amount of their income from

[5]Artists might initially be more concentrated in big cities such as New York and Los Angeles in which there are greater opportunities for success. In this case, the distributional concentration of concerts for Underdogs in big cities could be high. However, Underdogs' concerts would be dotted around the country at the same time. We focus on the trend of the distribution over the period.

[6] The website is a startup created by machine learning experts, and its comprehensive database has been created by using indexes of 135 different ticket vendors, venue websites, and local newspapers. It provides over 2 million historical concerts with place and date information

concerts in the U.S., and we suppose that this amount would be 10 million dollars in a year.[7] It is worth noting that only top 50~70 well-known artists could yield the greater than 10 million dollars of annual gross revenues in a year, and the list of these artists did not considerably change during our study period. This fact provides that this group of artists is indeed far more popular than others and capable of attracting audience in every place where they toured. On the other hand, considering the consistent appearance during the period of study, we selected artists who were ranked on the top 100 tour lists at least 4 times during the last decade.[8] By doing this, not only are artists who did not actively tour in the U.S. eliminated, but also we can sufficiently track the change of concert location distributions across years. According to these two criteria, 45 artists (e.g., Bon Jovi, Elton John, Madonna, and U2) are classified into the group of Superstars (Appendix A: A full list in the group of Superstars). The average career-beginning year of these artists is 1980 in the range from Willie Nelson (1956) to Rascal Flatts (1999), and the genres of artists vary from rock (e.g., Kid Rock, The Rolling Stones), pop (e.g. Britney Spears, Shakira) to country music (e.g., Toby Keith). They are no doubt top-notch artists historically in terms of successful careers, influence, and reputation in the music industry.

Second, it is important to clearly define Underdogs, because it is difficult to know all concerts performed in the U.S. during the last decade. For this reason, random sampling from the entire population is not an easy task, and it is required us to approach in a different way. To mitigate a possible sampling bias, we set the following sampling process for the group of Underdogs. First, we obtained information about all artists who held their concerts in the U.S. in 2007[9], and among these artists we exclude those who were listed in the annual top 100 tour ranking at least once between 2000 and 2009, because these artists were likely to be originally popular in larger regions. By doing this, 839 artists were selected. Second, in order to track the change of concert locations based on the artist fixed-effect, artists who performed their concerts in 2000 were chosen; as a result, 224 artists finally composed the group of Underdogs. In other words, these 224 artists kept their concert activities since 2000, and they could be regarded as relatively less known artists during the period of our study.

In summary, we examine the concert location distributions of both 45 Superstars and 224 Underdogs in the period from 2000 to 2011. The resulting sample includes 15,685 concerts by Superstars and 58,857 concerts by Underdogs. We then match each concert location to the geographic standard, CBSA.[10] We use this geographic standard for two reasons: first, following our research strategy outlines, it is required to establish geographic division appropriately. While either zip code- or county-level seems to be too specific, state-level is too broad. The present rules have defined 935 CBSAs in the U.S., and over 365 out of all CBSAs are classified as metropolitan areas in which the population is greater than 50,000. The population size across CBSAs shows a wide variation from nearly 20 millions in New York City area to 55,000 in Carlson City, NV. Second and more important, residents in each CBSA are socio-economically tied to the urban center by definition, and this fact indicates that people in an identical CBSA would tend to share infrastructure and facilities in close proximity. For instance, the largest CBSA, New York City area, covers regions of 24 counties in three states (New York, New Jersey, and Pennsylvania) in which most residents' routine lives relate to New York City.[11] Assuming that greater numbers of venues are located near the urban center, this division would make more sense.

---

[7] Since the average concert ticket price has been increasing during the last decade, the greater number of artists earned higher than 10 million dollars per year over the period. For instance, the annual gross revenues of 42 artists showed larger than 10 million dollars, whereas those of 75 artists showed over 10 million dollars.

[8] It is common that artists set tours after the release of new albums, and usually Superstars do not tour every year. If an artist generated high income from tours in a certain year and did not tour in the following years, it is difficult to observe the trend of concert locations over time. One example is that NSYNC was listed on the top in 2001 but they only performed concerts during that year. The other example is that a recently rising star, Justin Bieber, began his career in 2009 and his tour was listed on the top in 2009, but he did not have any history of concerts before this year.

[9] It would be better if we would have the list of all artists who performed their concerts in 2001, but we only have the information of 2007.

[10] CBSA is a U.S. geographic area defined by the Office of Management and Budget based an urban center of at least 10,000 people and adjacent areas that are socioeconomically tied by commuting.

[11] Refer to the Appendix B in which we depict a map of CBSA for New York City area.

Descriptive Statistics

Table 1 presents descriptive statistics of concerts and artists by every two year. The trend indicates that average numbers of concerts per active artists for both groups tend to increase over time; those for Superstars rose from 45.79 to 68.07, whereas those for Underdogs rose from 27.91 in 2000-2001 to 70.03 in 2010-2011.[12] This fact indicates that Underdogs' concerts have sharply increased over the period.

Table 1. Descriptive Statistics

Superstars (N=45)	Total Concert	Active artist	Concerts/artist
2000~2001	1,923	42	45.79
2002~2003	2,601	43	60.49
2004~2005	2,561	40	64.03
2006~2007	2,745	41	66.95
2008~2009	3,064	43	71.26
2010~2011	2,791	41	68.07
Total	15,685		
Mean	2,614	42	62.76
Std. Dev.	382	1	9.09
Underdogs (N=224)	Total Concert	Active artist	Concerts/artist
2000~2001	6,251	224	27.91
2002~2003	7,918	196	40.40
2004~2005	7,601	204	37.26
2006~2007	9,764	216	45.20
2008~2009	11,987	214	56.01
2010~2011	15,336	219	70.03
Total	58,857		
Mean	9,810	212	46.13
Std. Dev.	3,359	10	14.92

[12] For the group of Superstars, the number of concerts decreased in 2010-2011 from the previous period. This might be because of economic downturn which began in the late 2008.

Empirical Analysis and Results

Our empirical approach consists of two stages. First, we begin our analysis by presenting basic statistical properties and descriptive evidence of the distribution by using various cutoff points and calculating the Gini coefficients, etc. (Tan and Netessine, 2011). Second, in order to investigate distributional characteristics of Superstars and Underdogs, we empirically estimate salient trends and the level of concentration to the large regions with the introduction of a new measurement.

Basic Statistical Properties

The Pareto Principle is widely used to describe the skewness in distribution, and we first test the Pareto Principle by counting the number of concerts across areas. The result is reported in Figure 2: the horizontal axis represents data from the CBSA where the most concerts are held to the CBSA where the least concerts are held in a given time period, with 365 metropolitan CBSAs normalized between 0 and 100. The vertical axis denotes a fraction of cumulative number of concerts. Both graphs precisely show that 20% of top regions account for nearly 80% of concerts, indicating that the concert distributions follow the Pareto Principle. In other words, concerts by Superstars and Underdogs are both highly skewed to large cities.[13]

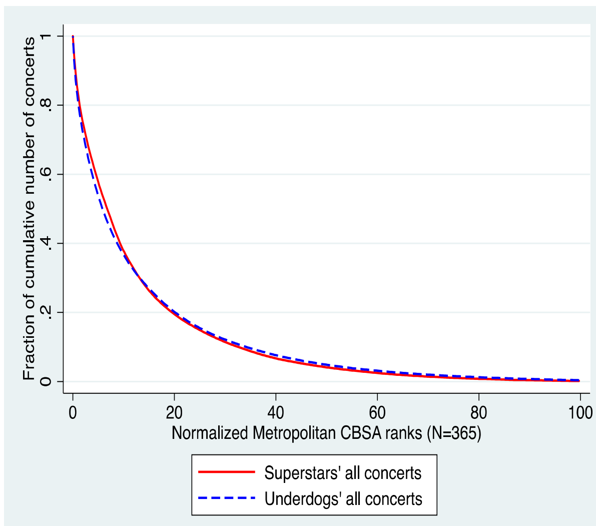

Figure 2. Skewness of Concert Distribution across CBSAs

Specifically, Table 2 provides the trend of concert distributions of two groups, Superstars and Underdogs. For convenience, we divide years to three periods by every four year. Corresponding the fact from Figure 3, the top 100 large areas (according to the size of population) out of 365 metropolitan CBSAs account for more than 80% of concerts. When we observe the trends of Superstars' concerts, the shares of large areas continued to decrease. For Underdogs' concerts, however, the shares of large areas (e.g., Top 3, Top 5, and Top 10) slightly increased over time. To illustrate this, the share of Top 5 CBSAs for Underdogs' concerts rose from 22.5% to 23.2%. In addition, CBSAs in which there was no concert decreased from 133 in 2000-2003 to 112 during 2008-2011 for Superstars and from 76 to 41 for Underdogs. All these observations support that both Superstars' and Underdogs' concerts were held in more dispersed places at a glance. Simultaneously, the concentration to the largest areas for Superstars concerts has been fallen, whereas that for Underdogs seemed to be slightly grown.

[13] To reflect this fact, we will take into account several cut-off points of highest-ranked CBSAs as well as a full set of CBSAs.

Table 2. Trends of Concert Distributions

All concerts	Superstars			Underdogs		
CBSA rank	2000-2003	2004-2007	2008-2011	2000-2003	2004-2007	2008-2011
Top 3	15.50%	15.35%	14.28%	16.36%	16.56%	16.72%
Top 5	21.21%	21.19%	19.60%	22.56%	22.58%	23.21%
Top 10	32.25%	30.69%	29.18%	33.84%	33.68%	33.65%
Top 20	48.26%	44.48%	44.42%	48.83%	49.48%	48.43%
Top 30	60.05%	55.60%	56.54%	59.36%	59.38%	58.54%
Top 50	74.90%	71.04%	72.06%	71.36%	71.70%	70.70%
Top 100	89.57%	87.60%	87.97%	86.54%	86.94%	85.88%
<100	10.43%	12.40%	12.03%	13.46%	13.06%	14.12%
	Superstars			Underdogs		
CBSA (N=365)	2000-2003	2004-2007	2008-2011	2000-2003	2004-2007	2008-2011
Concert=0	133	128	112	76	58	41
Concert>0	232	237	253	289	307	324
Concert>5	115	133	133	190	201	234
Concert>10	73	100	97	152	165	205
Concert>30	33	38	41	79	85	108
Concert>50	26	33	37	62	80	96
Concert>100	6	5	8	36	42	64

We then use the Gini coefficient and the Lorenz curve to address the concentration of concerts in each group over time by applying two additional cutoff points (Top 50 and Top 100 CBSAs). In addition to the measurement on concert-level, we also conduct artist-level analysis here. The probability of artist's visit to a particular CBSA is computed by the number of artist who performed a concert in CBSA divided by the total number of active artists in the given time.[14] The result is reported in Table 3. All measured Gini coefficients for Superstars declines over the period and in each cutoff case, suggesting that the decreasing trend is more salient in the case of Top CBSAs. This fact indicates that the Superstars' concerts became less concentrated in cases of Top 100 or Top 50 CBSAs rather than all CBSAs, which could be inferred that in any time periods Superstars are less likely to visit too small regions where the population size is less than 500,000. Moreover, according the Gini coefficients of artists, the share of visits to big cities for Superstars has become reduced, suggesting that Superstars relatively toured fewer times in big cities.[15] On the other hand, the concentration level of concerts for Underdogs slightly rose. Despite the fact that the change seemed not to be considerable, the directions of the trend between Superstars and Underdogs are fairly different. Gini coefficients of concerts and artists have increased from 0.789 and 0.808 to 0.724 and 0.739, respectively.

[14] For instance, when the total number of active Underdog artists was 224 in 2000-2001 and 150 artists performed their concerts in New York City, the probability of visit New York City will be 150/224.

[15] Because the total number of concerts tended to rise over time, this does not mean that Superstars performed fewer in those big cities in an absolute term. In other words, the big cities would be the most places for them, but they have begun touring more in smaller cities.

Table 3. Trend of Gini Coefficient						
# of Concerts	Superstars			Underdogs		
All Concerts	2000-2003	2004-2007	2008-2011	2000-2003	2004-2007	2008-2011
All CBSAs	0.735	0.710	0.691	0.789	0.801	0.808
Top 100 CBSAs	0.567	0.533	0.491	0.575	0.576	0.572
Top 50 CBSAs	0.388	0.372	0.316	0.412	0.415	0.410
# of Artists	Superstars			Underdogs		
All Concerts	2000-2003	2004-2007	2008-2011	2000-2003	2004-2007	2008-2011
All CBSAs	0.646	0.627	0.625	0.724	0.742	0.739
Top 100 CBSAs	0.441	0.408	0.388	0.458	0.462	0.431
Top 50 CBSAs	0.228	0.214	0.195	0.266	0.270	0.234

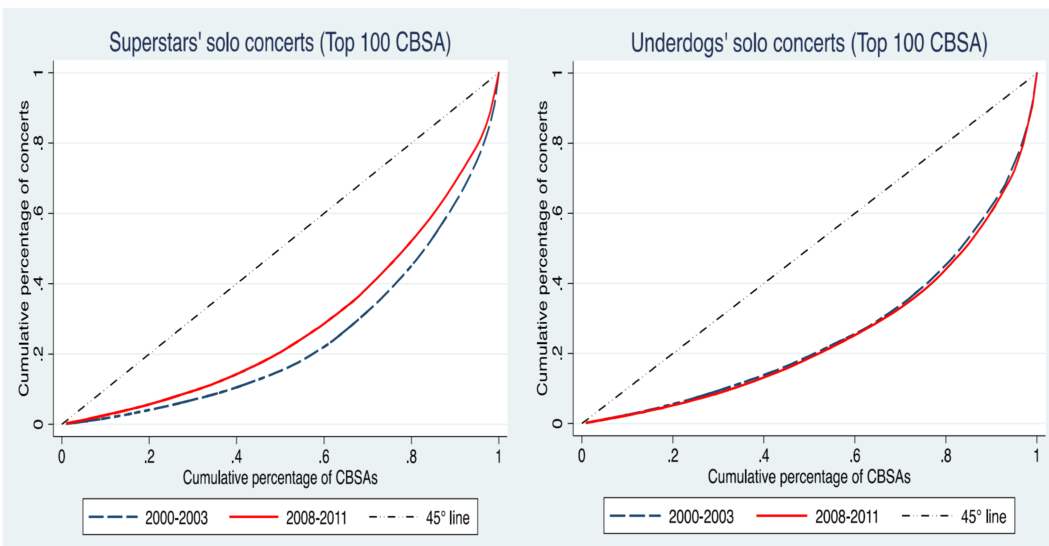

Figure 3. Lorenz Curve in Top 100 CBSA

Figure 3 visually illustrates the comparison of the Lorenz curve between the early period (2000-2003) and the late period (2008-2011). Similar to the findings in Table 3, the Lorenz curve in the late period for Superstars markedly lies above that in the early period, whereas, for Underdogs, the Lorenz curve in the late period slightly lies below that in the early period. However, these tools do not allow us to conclude whether or not such a difference is statistically significant (Brynjolfsson et al., 2011). We thus fit the concert and rank data to the log-linear relationship and compare the coefficient obtained to examine the concentration of the distribution. On the top of that, we conjecture that counting only the number of concerts may cause a nontrivial bias in examining the actual behaviors of artists. Accordingly, we introduce a new measurement to explain the concentration ratio more accurately.

The Necessity to Introduce a New Measurement

The series of analyses and observations in the previous section demonstrated several important distributional dynamics by using concert shares across region and time. However, this approach is not sufficient to measure a proper concentration ratio, because we did not account for the frequency of artist's visits in company with the concert share. To illustrate this, suppose that there are 100 concerts and 10 artists in a given period.

In case I, even though a small share of concerts (30%) is held in Location B, 8 out of 10 active artists performed in this venue. In this case, it is not reasonable to say that the share of Location A in this concert market is 70% without accounting for how many artists performed their concerts. Case II shows a more extreme example: even if the share of concerts is identical between Location C and Location D, all artists perform in Location C and only half of the artists visit Location D. That is, considering only the concert shares (50% vs. 50%), may cause a significant bias to present actual concentration ratio. This particular attribute of the concert market drives us to introduce a new measurement by which both concerts' and artists' share would have to be considered in a comprehensive way. We thus introduce a new index, Concentration Propensity (or CP) to measure the propensity of concentration by taking into account both concerts and artists.

$$Concentration\ Propensity_{it}(or\ CP_{it}) = \frac{Concert_{it}}{\sum_{j=1}^{J} Concert_{jt}} \times \frac{Artist_{it}}{\sum_{m=1}^{M} (Artist_{mt} = 1 \text{ if } Artist\ m \text{ is active at time } t)} \quad (7)$$

where i is CBSA, t is year. J and M denote the total number of CBSAs and the total number of artists, respectively. Specifically, the first term defines a simple normalization, which is a share of concerts in CBSA i at year t, and the second term specifies how many artists visit CBSA i at year t out of total active artists at year t.[16] By doing this, we can control the possibility of multiple concerts by a single artist in a certain area (e.g., It is common to plan to perform concerts more than once in large cities such as New York City).[17]

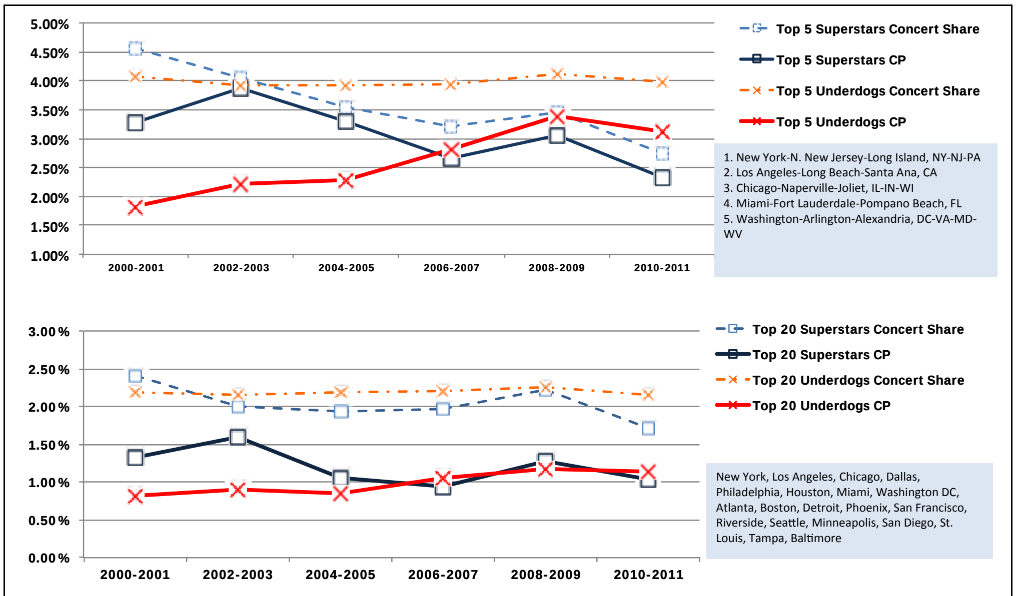

Figure 4. Trend of Concert Shares in Top 5 and Top 20 CBSAs

We also see how the suggested CP ratio is different from the concert ratio, and Figure 4 depicts the trends. Presenting cases of Top 5 CBSAs and Top 20 CBSAs, the solid line and the dot line denote the CP ratio and the concert share, respectively. While the concert share of Top 5 CBSAs for Superstars tend to decrease in the first

[16] For instance, total number of active Superstar artists in 2010-2011 is 41, and this is a denominator of the second term. In the same period, 35 artists visited New York area, 18 artists visited Pittsburgh, PA, whereas only 6 artists visited Tucson, AZ.

[17] By using this new measurement, the discrepancy between Location A and Location B in the earlier Case I is reduced (i.e., Location A's CP is 0.7*0.5=0.35 and Location B's CP is 0.3*0.8=0.24). In Case II, the CPs of Location C and Location D is different (0.5 vs. 0.25) unlike the same concert shares.

graph, the CP of the Top 5 appears to be relatively flat. This fact indicates that higher initial concert share came from the small number of active artists, and this value seemed to be over-specified without accounting for the shares of artists in these areas. The case of Underdogs is more striking. Underdogs' CP ratio of the Top 5 CBSA rises sharply over time, whereas the concert shares seem to be stable at 4%. This fact implies that the larger shares of Underdog artists have become performed in these big cities over time. The second graph in Figure 4 specifies analogous trends in the case of the top 20 CBSAs. All this evidence supports the necessity of introducing a new measurement, because the normalization of concert shares is neither sufficient to present the actual concentration tendency nor possible to bear a bias.

Empirical Analysis and Results

We now delve more deeply into the empirical properties of geographic concert distribution, focusing on how the concert distributions have evolved. In other words, the purpose of this section is to demonstrate the observable implications of the dynamics discussed above. We thus employ a log-linear relationship to test the Long Tail trend based on the fact that the size of organization *S* is systematically related to its rank *R* according to the Pareto law $S \times R^{\beta} = A$, where β and A are constants. In our data, CBSAs are naturally ranked each year by the suggested Concentration Propensity (*CP*) ratio, and the Pareto curve implies the following log-linear relationship between CP and rank:

$$\log(CP_{it}) = \beta_0 + \beta_1 \log(Rank_{it}) \qquad (8)$$

where *i* and *t* denote CBSA and year, respectively. Previous research suggests that a distinguishing feature of power law is a straight line in the log-log plot. This has been used to describe the distribution of city size (Zipf, 1949), and given our specification, β_1 measures how quickly the CP level goes down as the rank increases. If the distribution has a longer tail (or is less concentrated), then β_1 would be lower in absolute value (Brynjolfsson et al., 2011), indicating that live tours are more dispersed and that the share of large cities decreases.

We first estimate Equation (8) separately for the two groups (Superstars and Underdogs) across periods with all samples, and we cut off the data by Top 100 CBSAs.[18] Table 4 presents the result from Superstars. All coefficients are highly significant, and higher R^2 values suggest that the log-linear relationship fit well. In Columns (1) to (6) with all CBSAs in Table 4, coefficient estimates of β_1 tend to reduce over the period, but the discrepancy is not substantial. However, considering the Top 100 CBSA, the estimated coefficient sharply increased from -1.74 in 2000-2001 to -1.28 in 2010-2011. Estimated constant also significantly decreased from -0.32 in Column (1) (2000-2001) to -1.37 in Column (6) (2000-2011). In other words, the Long Tail trend is observed from the Top 100 CBSA, indicating that Superstars relatively visited more middle-sized regions to perform more concerts instead of focusing on the large cities, which supports our hypothesis 2. This finding can be reasonably interpreted that Superstars' behaviors in the concert location choice have been changed within Top 100 CBSAs, because they are less likely to go to small cities even in the later period.

The result from the case of Underdogs is reported in Table 5. As opposed to the case of Superstars, the coefficient estimates in all CBSAs significantly decreased over time, and the constant has sharply become growing. β_1 has dropped from -2.79 in Column (1) (2000-2001) to -3.01 in Column(6) (2010-2011), and β_0 has grown from 2.49 in 2000-2011 to 4.38 in 2010-2011. This is evidence of the strengthened Power Law trend over the period, which supports Hypothesis 3. However, there was not a significant shift in the analysis with Top 100 CBSA. The result can be interpreted in a different way from the Superstars' case. Underdogs were more likely to perform their concerts in a wider range of regions (e.g. beyond Top 100 CBSA) in the early periods; therefore, their changed behaviors of concert location choice may affect the distributional shift of the 'All CBSAs' case rather than of the 'Top 100 CBSA' case.

[18] The reason that we take this Top 100 CBSA as an important cut-off is that Figure 3 and Table 2 illustrated the more than 85% of concerts were held in these regions. Considering Top 100 CBSAs, we can observe the dynamics within these places.

Table 4. Long Tail vs. Power Law Estimates (Rank-Size): Superstars

All CBSAs	(1) 2000-2001	(2) 2002-2003	(3) 2004-2005	(4) 2006-2007	(5) 2008-2009	(6) 2010-2011
log(rank)	-2.369***	-2.469***	-2.305***	-2.380***	-2.364***	-2.308
	(0.055)	(0.055)	(0.054)	(0.047)	(0.062)	(0.052)
Constant	1.588***	2.385***	1.863***	1.703***	2.062	1.939
	(0.248)	(0.257)	(0.251)	(0.226)	(0.286)	(0.248)
Adjusted R squared	0.897	0.890	0.879	0.892	0.859	0.876
Sample Size	208	233	249	277	243	275
Top 100 CBSAs	(1) 2000-2001	(2) 2002-2003	(3) 2004-2005	(4) 2006-2007	(5) 2008-2009	(6) 2010-2011
log(rank)	-1.740***	-1.569***	-1.406***	-1.338***	-1.373***	-1.281***
	(0.075)	(0.065)	(0.052)	(0.042)	(0.055)	(0.056)
Constant	-0.320	-0.461*	-1.021***	-1.343***	-1.077***	-1.375***
	(0.282)	(0.244)	(0.198)	(0.159)	(0.209)	(0.211)
Adjusted R squared	0.845	0.856	0.877	0.909	0.861	0.841
Sample Size	100	100	100	100	100	100

Note: Standard errors in parentheses. *** $p<0.01$; ** $p<0.05$; * $p<0.1$.

Table 5. Long Tail vs. Power Law Estimates (Rank-Size): Underdogs

All CBSAs	(1) 2000-2001	(2) 2002-2003	(3) 2004-2005	(4) 2006-2007	(5) 2008-2009	(6) 2010-2011
log(rank)	-2.796***	-2.838***	-2.843***	-3.004***	-3.035***	-3.014***
	(0.050)	(0.049)	(0.048)	(0.047)	(0.054)	(0.047)
Constant	2.494***	3.065***	2.987***	3.808***	4.105***	4.384***
	(0.247)	(0.248)	(0.243)	(0.242)	(0.275)	(0.250)
Adjusted R squared	0.903	0.896	0.903	0.909	0.889	0.895
Sample Size	331	387	370	403	391	468
Top 100 CBSAs	(1) 2000-2001	(2) 2002-2003	(3) 2004-2005	(4) 2006-2007	(5) 2008-2009	(6) 2010-2011
log(rank)	-1.583***	-1.554***	-1.517***	-1.565***	-1.559***	-1.449***
	(0.042)	(0.042)	(0.038)	(0.045)	(0.046)	(0.043)
Constant	-1.575***	-1.414***	-1.530***	-1.140***	-1.041***	-1.214***
	(0.160)	(0.160)	(0.145)	(0.172)	(0.173)	(0.162)
Adjusted R squared	0.933	0.930	0.939	0.922	0.920	0.919
Sample Size	100	100	100	100	100	100

Note: Standard errors in parentheses. *** $p<0.01$; ** $p<0.05$; * $p<0.1$.

Discussion and Conclusion

The decision to choose a concert location is an activity in which profit maximization and social influences play a key role. The touring location distribution is an outcome of the decision that combines prediction of potential incomes. As researchers previously examined, the proliferation of online channels has shifted revenue structures of the music artists (Oberholzer-Gee and Strumpf, 2009). Concerts *per se* have become more important for artists under the circumstance in which file sharing is widespread and online viral marketing is a nontrivial part for artists. In this study, we provide theoretical and empirical evidence to shed additional light from a different perspective on the existence of the Long Tail effect in the Internet age. The original definition of the Long Tail effect by Anderson (2006) focused on the demand side (i.e., niches will constitute a larger proportion of demand on the Internet). The concert industry used to be dominated by a few Superstars, but advanced technology allows lesser-known artists to perform more concerts, which is consistent neither to the "Superstar effect" suggested by Rosen (1981) nor to the "winner-take-all" theory. Our finding indicates that artists have performed more concerts over time, and this fact indeed allows consumers to experience more concerts than before (i.e., product variety). On the top of that, this finding accords with the original definition of the Long Tail effect from the perspective of the demand. We then reverse the lens to the supply side, and our analysis suggests that concert distribution and the level of concentration to the large cities is significantly different between Superstars and Underdogs over the last decade. By evaluating the distribution of concerts from several angles, we find evidence that Superstars' concert distribution follows the Long Tail effect, whereas Underdogs' distribution follows the strengthened Pareto curve over the period in terms of geographic location. In this context, which factors are associated with changing trends in concert location distributions? We speculate that the growing Internet adoption is correlated to the dynamics. Cho et al. (2012) showed that higher broadband penetration rate is associated with the higher number of concert attractions, and the impact is greater for Underdogs than for Superstars and for smaller cities than for larger cities. According to Chris Anderson who coined the term, Long Tail, for the first time, he specified that liver performance is the fast growing part of the music industry and the concert market represents the case of the Long Tail effect.[19]

From the perspective of industrial organization, the concert industry is seen as a monopolistic competitive market that controls the process of the creation of songs, music sales, and live tours. A concert is a non-digitized experience good whose value can be fully recognized only when it is consumed. While recorded music can be substituted for newly developed digital formats such as MP3, the unique experience in a concert venue cannot be replaced by watching video clips of previously performed concerts. In this respect, Internet adoption could be beneficial for touring artists in several aspects. From the perspective of demand, consumers may obtain more information about artists and upcoming tours in their town by using online search tools with lower costs. Not only can they book a ticket through online channels (e.g., Ticketmaster.com), but more direct interactions and communications with artists or other fans may also lead to high likelihood of attending concerts. It is not difficult to observe that artists promote their upcoming concerts via Twitter, Facebook, and YouTube channels. Our empirical work implies that lesser-known artists utilize this opportunity and tour more frequently in large cities.

On the other hand, from the perspective of supply, artists would need to perform more concerts in the era of file sharing in order to recoup declining recording revenues. Since well-known artists would lose a larger portion of income, this supply-side trigger is more likely to be associated with Superstars. When they increase the number of concerts, our finding suggests that they tour in smaller cities where they had not performed in the past. This phenomenon accords with the Long Tail effect in terms of geographic distribution. Our analysis did not cover worldwide tours outside of the U.S., but anecdotal cases indicate that there have been a greater number of global tours by Superstars over the period.

Our paper is not without limitations. Due to the limited available data, we count only the number of concerts in order to examine its distribution. As an extension to this study, more promising research can be conducted with data regarding concert ticket price and the number of attendees per concert. Another interesting future venue for relevant research would be to compare artists who proactively utilize social media to artists who do not in the period after the emergence of social networking sites. In any event, our findings suggest a new layer of study for the relationships between the Internet and the music industry from a broader point of view.

---

[19] Source: http://www.longtail.com/the_long_tail/2007/01/give_away_the_m.html

References

Anderson, C. 2006. The long tail: Why the future of business is selling less of more, New York: Hyperion.

Becker, S. G. 1965. "A Theory of the Allocation of Time," The Economic Journal (75:299), pp. 493-517.

Blackburn, D. 2007. "The Heterogeneous Effects of Copying: The Case of Recorded Music," working paper, Harvard University, Cambridge, MA.

Boorstin, E.S. 2004. "Music Sales in the Age of File Sharing," Senior Thesis, Princeton University.

Brynjolfsson, E., Hu. Y. J., and Smith, M.D. 2006. "From Niches to Riches: The Anatomy of the Long Tail," MIT Sloan Management Review (47:4), pp. 67-71.

Brynjolfsson, E., Hu. Y. J., and Smith, M.D. 2010. "Long Tails vs. superstars: The effects of information technology on product variety and sales concentration patterns," Information Systems Research (21:4), pp. 736-747.

Brynjolfsson, E., Hu. Y. J., 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), pp. 1373-1386.

Cho, D., Telang, R., and Smith, M.D. 2012. "Live Show Everywhere: Distribution Dynamics and Internet Influence on Concert Location," Proceedings of the 2012 45th Hawaii International Conference on System Sciences (HICSS), pp. 4495-4504.

Connolly, M., and Krueger, A. B. 2006. "Chapter 20 Rockonomics: The Economics of Popular Music," in Handbook of the Economics of Art and Culture, V.A. Ginsburg and D. Throsby (eds.), Elsevier, pp.667-719.

Dejean, S. 2009. "What Can We Learn from Empirical Studies About Piracy?" CESifo Economic Studies (55:2), pp. 326-352.

Elberse, A., and Oberholzer-Gee, F. 2007. "Superstars and Underdogs: An examination of the long tail phenomenon in video sales," Marketing Science Institute (4), pp. 49-72.

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

Hong, S-H. 2005. "Measuring the Effect of Digital Technology on the Sales of Copyrighted Goods: Evidence from Napster," working paper, University of Illinois at Urbana-Champaign.

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

Michel, N. 2006. "The Impact o Digital File Sharing on the Music Industry: An Empirical Analysis," B.E. Journal in Economic Analysis and Policy: Topics in Economic Analysis and Policy (6:1), pp. 1-22.

Mortimer, J. H., Nosko, C., and Sorensen, A. 2012. "Supply responses to digital distribution: Recorded music and live performances," Information Economics and Policy (24:1), pp. 3-14.

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

Oberholzer-Gee, F., and Strumpf, K. 2009. "The File Sharing and Copyright," Harvard Business School working paper.

Peitz, M., and Waelbroeck, P. 2004. "The Effect of Internet Piracy on Music Sales: Cross-section Evidence," Review of Economic Research on Copyright Issues (1:2), pp. 71-79.

Rosen, S. 1981. "The Economics of Superstars," The American Economic Review (71:5), pp. 845-858.

Smith, M., and Telang, R. 2009. "Competing with Free: The Impact of Movie Broadcasts on DVD Sales and Internet Piracy," MIS Quarterly (33:2), pp. 312-338.

Tan, T. F. and Netessine, S. 2011. "Is Tom Cruise Threatened? An Empirical Study of the Impact of the Product Variety on Demand concentration," in Proceedings of the International Conference on Information Systems, pp. 1-18.

Zentner, A. 2006. "Measuring the Effect of File Sharing on Music Purchases," Journal of Law and Economics (49:1), pp. 63-90.

Zipf, G. 1949. Human Behavior and the Principle of Least Effort. Addison-Wesley, Cambridge, MA.