DISSERTATION

THE CREATIVE DESTRUCTION OF THE "WINNER-TAKE-ALL" SOCIETY? PROPERTY RIGHTS AND THE ECONOMICS OF THE LONG TAIL IN THE MUSIC INDUSTRY

Submitted by

Ian Strachan

Department of Economics

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2010

Doctoral Committee:

Department Chair: Steven J. Shulman

Advisor: Ronnie J. Phillips

Gregory D. Graff
Nancy Jianakoplos
David Mushinski

ABSTRACT

THE CREATIVE DESTRUCTION OF THE "WINNER-TAKE-ALL" SOCIETY? PROPERTY RIGHTS AND THE ECONOMICS OF THE LONG TAIL IN THE MUSIC INDUSTRY

Technological change has always impacted the music industry, which is now absorbing the destructive effects of the digital revolution. The Internet and MP3s have allowed for illegal downloading and file sharing. For producers, revenue streams have run dry because consumers can get their music for free. The old property rights regime has been eroded for incumbents and a market failure is imminent. But the digital revolution also has the potential to create opportunity for entrepreneurial artists and firms who are able to utilize new technology for disseminating their content. This dissertation utilizes a random sample of recorded music unit sales for 2,051 artists from 2004 to 2008. The data is used to test if the digital revolution has created a 'long tail effect' where less popular and nascent artists enjoy more sales, or a 'superstar effect' where a small number of top artists take the lion's share of sales. I find that the market is characterized by an extremely skewed sales distribution profile which reaches a peak in both sales and inequality in 2006 while sales and inequality decline thereafter. I also find a superstar effect in digital formats and a long tail

effect in non-digital formats across all five years. The 'middle class' of artists also declines steadily. For property rights, these changes in the sales distribution profile highlight the importance of retaining excludability through bundling content together. Bundling can still be facilitated by copyright collectives and intermediaries.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

I. Introduction

For most of the Twentieth Century, the established business model in the music industry was one where commercial success necessitated that artists sign contracts with major record companies. The structure of the industry was what Frank and Cook (1995) called a 'winner-take-all' market. There was a two-tiered structure where the top artists earned enormous incomes and sold millions of records, while the bulk of artists, most of whom were not signed to major companies, were left to divide the remainder of the revenue pie. For example, because people are willing to pay much more to see Mick Jagger or Paul McCartney perform, or prefer their recordings to those of other artists, people have less to spend on live performances or recordings of other artists. Building on the work of Rosen (1981), Frank and Cook (1995) argued that the development of a 'winner-take-all society' threatened to exacerbate income inequality and misallocate resources as more people tried to achieve these high incomes but fail.

At the beginning of the 21st century, a confluence of new digital technology became the death knell for the established recorded music industry. Creative destruction has made record companies' business models irrelevant. Digital technology has had an adverse effect on other industries as well, such as newspapers. When it comes to music and newspapers, consumers now have more choice than ever before and can often access what they want for free. On the supply side, creators, musicians, retailers, and publishers potentially benefit

from digital technology because most production and distribution processes are streamlined and more efficient. Unfortunately, their revenue streams have dried up in the process because scarcity has vanished. This suggests a market failure is present. There is uncertainty about how the market will adapt to the new digital paradigm.

A. Statement of the Problem: Digital Creative Destruction

Today, the Internet and digital technology have revolutionized the music industry. This has been destructive to old business models. But many artist entrepreneurs have created different ways to earn their living in the music business. Signing with a major record company is no longer needed when the costs of recording and distributing technology have been dramatically reduced. New contractual arrangements and business models are being used by both established and new artists. The fundamental research question addressed in this dissertation is whether these recent changes in technology coupled with entrepreneurial initiative in the form of new business models can change the 'winner-take-all' structure and create a 'new artistic middle class' that will better distribute the revenues generated by the industry.

B. Research Questions

Though live performance retains the characteristics and problems of 'winner-take-all-markets', in the past decade, the development of digital technology has dramatically lowered the cost of producing and distributing recorded music. A lower cost structure suggests an

increase in supply, which enables consumers to buy more as prices have fallen. But this shift also offers the possibility of dramatically increasing the variety of digital recordings available as barriers to entry are lowered. This expansion of availability has been labeled 'the long tail' (Anderson 2006). Consumers often first learn of a performer through digital recordings available online through streaming and recommendation services such as Pandora or Rhapsody. Online retailers can also afford to stock more choices. Because of this, the potential market for nascent and niche artists is expanded as consumers can more easily find new music. Though this change in supply may enhance the chances of nascent artists becoming 'superstars', it also means an increased supply of rare or obscure recordings of superstars.

The questions to be answered are the following:

1) Is there a 'superstar' effect in the market for recorded music today and is it significantly different from the past?

2) Is there a 'long-tail' effect in the market for recorded music today and is it significantly different from the past?

3) Is there some combination of the 'superstar' and 'long-tail' effects?

C. Hypotheses

Much of recorded music is no longer phonograms – in physical form. It is now digital data, where production, distribution, inventory, and reproduction costs are approaching zero. This may be represented as an outward shift of the industry supply curve. Prices are falling not only because costs are going down, but also because consumers do not need to purchase music anymore. This may be represented as the demand curve for

recorded music shifting to the left. Before the digital revolution, there were only physical records, which created rivalness and excludability. Poor-quality copies created excludability. Now, most recorded music is available for free. With the right technology, consumers can reproduce and download recorded music for free without compromising the recording's quality in a meaningful way. This change creates clear disadvantages for many producers of music. But for consumers, digital music has lowered prices, increased choice, and created new modes of consumption.

There is clear evidence of an increase in supply, but does digital technology affect demand as well by expanding the types of music consumers like? Most of the uncertainty for industry operators is about what will happen on the demand side and to the resulting distribution of sales. This implies existing business models of incumbent artists and firms could be replaced by new business models of entrepreneurial artists and firms who are able to better meet demand changes and capture revenue streams from those changes.

Chris Anderson, former chief editor of 'Wired' magazine, wrote a widely influential article in 2004 and a subsequent book in 2006 entitled 'The Long Tail: Why the Future of Business is Selling Less of More'. The economics of the long tail refers to the view that the rise of Internet commerce, online retailers, and digitization will bring about a new economic reality. Hits will matter less and niche markets – the 'long tail' of the sales distribution – will matter more as a proportion of total sales. This is illustrated by Anderson in Figure 1. Digitization applies to different contexts and means the transformation from analog to digital form, and most importantly here, data on a computer. Decreased distribution costs and inventory carrying costs imply increased supply and increased product variety. In response, consumption patterns – demand – will diversify because preferences are better matched, and sales will move away from a small number of high-volume hits towards a

larger number of low-volume niche products. This larger number of less-popular products will make up a relatively larger proportion of revenues for retailers.

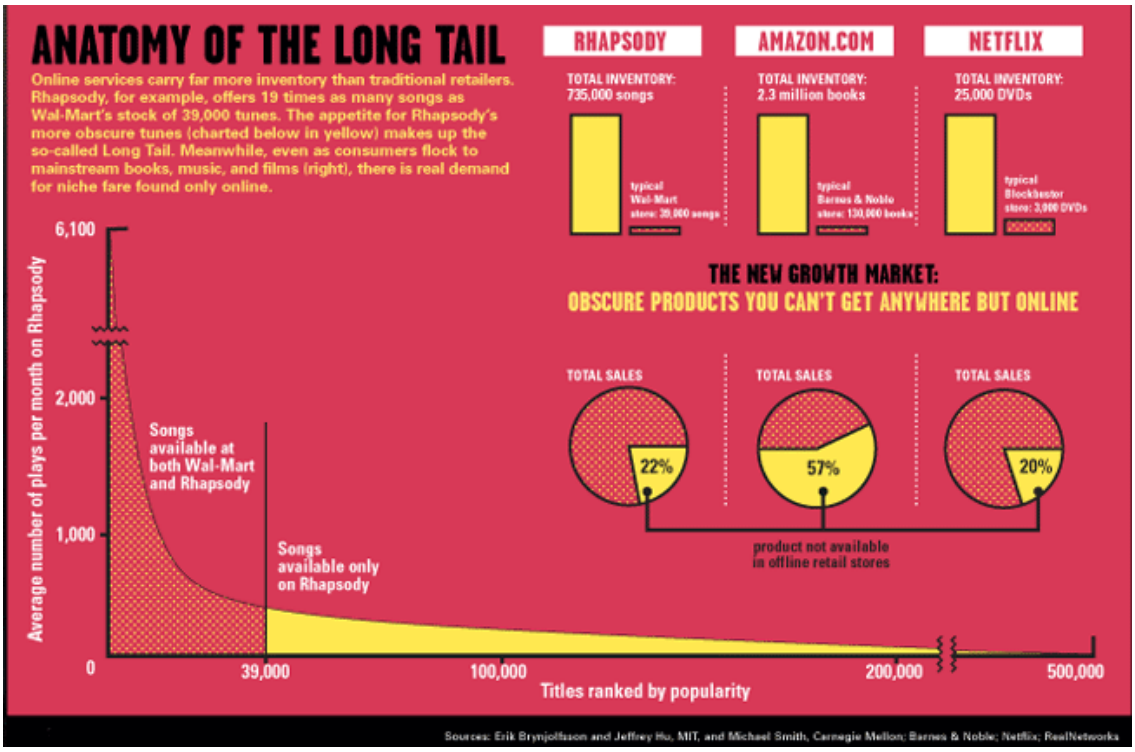

Figure 1: Anatomy of the Long Tail.

This is from Anderson's (2004) original article and touts the relative importance of the 'tail' of the sales distribution as compared to the 'head' of the sales distribution. This is justified by the comparison of limited brick-and-mortar store offerings compared to digital platforms which have non-rival and more diverse inventories.

In addition, consumer search costs are lowered by online search, recommendation, and filtering tools (e.g. Brynjolfsson, Hu, and Simester 2006). This potentially helps less-known and new artists gain exposure. The earlier work of Joseph Schumpeter explains that the entrepreneurs who break up the economic structure are new, small firms coming from the outside (Schumpeter 1911, cited in Audretsch 2008). This is observed with digital technology as incumbents in the recorded music industry have seen devastating sales declines. Entrepreneurial artists and other creators could leverage this situation of creative destruction. According to the long tail thesis, the democratization of music production,

distribution, and consumption should shift the music industry's balance of power away from incumbents and in favor of these entrepreneurs.

If the long tail theory supposes that consumer preferences are better matched by increased variety, then consumers must have diverse tastes, or at least eventually they discover these diverse tastes. This may be a tenuous assumption. The superstar hypothesis, on the other hand, predicts consolidation of preferences on the most popular products. Low popularity products do not sell as much as a dominant choice in the mass market. This choice gets stronger through shared consumer preferences and becomes more dominant and then sells even more. Imperfect substitution assures this effect (Rosen 1981). Because reproduction, i.e. the marginal cost, of a recording is cheap, what is popular becomes excessively profitable. Therefore, the rise of Internet commerce, online retailers, and digitization should enhance the superstar effect and the winner-take-all model should prevail. So, the decline in barriers to entry into larger markets can benefit top performers and well-capitalized firms disproportionately (Frank and Cook 45).

Given the changes in technology and institutions in the music industry, what will the music industry look like in the future? Will the 'winner-take-all' structure enhance superstars or will there be an improved distribution of industry revenues? This dissertation evaluates three hypotheses about the emerging structure of the industry.

Hypothesis 1. *There is a 'superstar' effect in the market for recorded music today. The superstar hypothesis* is that sales have become relatively more skewed toward hits over time. If the distribution of recorded music sales has shifted toward top quantiles over time at the expense of all the other quantiles, this hypothesis will be supported (illustrated in Figure 2).

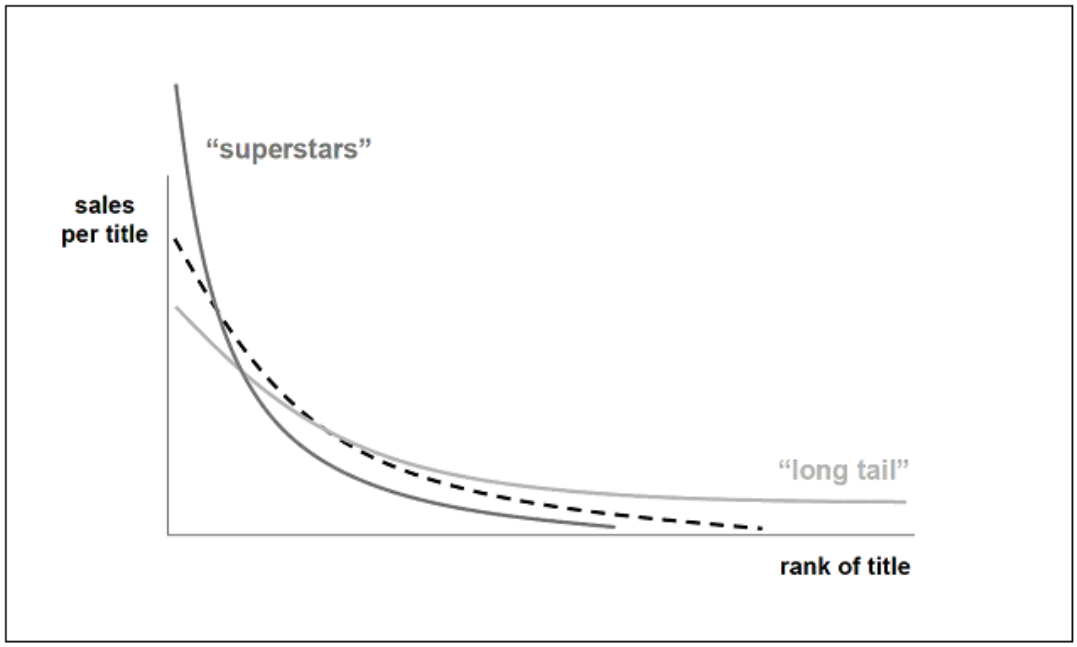

Figure 2: Illustration of the Superstar and Long Tail Effects.
This is from Elberse and Oberholzer-Gee (2008) who examine both hypotheses on sales data for movies, another industry affected by the digital revolution.

Hypothesis 2. *There is a 'long-tail' effect in the market for recorded music today that helps new or insurgent artists increase sales at the expense of established or incumbent artists. The long tail hypothesis* is that sales have become relatively less skewed toward hits and superstars over time. If the distribution of recorded music sales has shifted toward bottom quantiles over time at the expense of all other quantiles, this hypothesis will be supported (illustrated in both Figures 1 and 2).

The superstars hypothesis predicts that demand will become more homogeneous. The long tail hypothesis predicts that demand will become more heterogeneous. Paradoxically, the same technology change which leads to the leveling of the playing field in the long tail thesis could lead to the superstar thesis. Many researchers with relevant data believe the popular long tail theory has recently become dogmatic: nothing more than 'web utopian fantasies' (e.g. Gomes 2006).

Hypothesis 3. *There is some combination of the 'superstar' and 'long-tail' effects. The hybrid hypothesis* is a combination of the effects of the above two hypotheses. If the distribution of recorded music sales has shifted such that both top and bottom quantiles increase at the expense of the middle quantiles, this hypothesis will be supported.

This hypothesis could also be referred to as the 'shrinking middle class' hypothesis. Some recent studies have found a distribution becoming skewed toward the tail and the head simultaneously while the distribution is tucked towards the origin as absolute sales fall across the entire distribution, but relatively more for the middle quantile (Elberse and Oberholzer-Gee 2008, and Page and Garland 2009). This 'hit-heavy, skinny tail' implies greater inequality. However, unlike the superstars hypothesis, superstars in this distribution see a shrinking absolute share, even though their relative share is increasing. But this is also true of the observations in the tail – the 'underdogs' – because the tail becomes longer but thinner.

Evaluation of these hypotheses is of great importance for the future of the music industry. The long tail theory itself became a hit among many industry operators and observers because it fit many peoples' world view of increased competition and diversity. Some skepticism followed as creative industries continued to see shrinking sales. If there is a long tail effect, a leveling of the playing field is occurring, likely benefiting new entrepreneurs. If there is a superstar effect, consolidation by the few at the top of the hierarchy is occurring, likely benefiting incumbents. As Page and Garland (2009) ask, if indeed the tail of available niche products lengthens – a supply side effect – will it then 'fatten' with sales – a demand-side effect? This will depend upon how the demand for recorded music changes in response to the digital revolution (Tschmuck 2006). A change in the distribution of revenues from incumbents to entrepreneurs would imply increased creative freedom and innovation. On the other hand, investment in new talent could be dampened due to the incumbent firms'

heavy losses. This latter prediction is suggested by Schumpeter's later work (Schumpeter 1942, cited in Audretsch 2008). Innovation needs to be funded and large, incumbent firms can make the necessary investments. If startup costs are prohibitive to get a new creative act up and running, well-capitalized record labels may be necessary. But, both effects could cause innovation in business models as artists and firms cooperate or defect in enforcing copyrights, or shift to different approaches to making money.

To examine these questions, this dissertation quantifies year-end sales in the recorded music industry over a five year period from 2004 to 2008. Specifically, a random sample of 2,051 artists, 7,010 titles – which are albums, singles and videos – included on 1,836 labels are observed to see if and to what degree sales distribution profiles are changing from year to year. The changes are measured by a variety of techniques also used in other studies of this nature. But, no other research has addressed all of the above hypotheses in the recorded music industry with such a large sample or in such a recent, relevant time frame.

Many retailers have taken Anderson's long tail theory to heart and implemented it in practice. This allows their customers more choice, but also increases costs which make this strategy risky if consumers do not support the vast array of niches available. As Will Page, chief economist for a copyright collective in the U.K., observes, "now we've seen what happens when tens of millions of choices are thrown in the air and land on the floor" (Orlowski 2008). Shedding light on this debate is significant for the strategies of content creators, artists, intermediaries, and retailers. In the old technological paradigm, in the brick-and-mortar world of rivalness in inventory, there were limits to the products which were given shelf space. Those products which were worthy of marketing expenditure were chosen very carefully. Content creators, artists, intermediaries, and retailers now must determine whether or not to change their marketing and inventory strategies.

This debate applies to globalization as well. Many view technology changes and globalization in a positive sense, as with the 'web utopian' ideas of Anderson where nascent creators flourish. But still others view the same technology changes and globalization in a negative sense, predicting that culture will be homogenized. The Internet potentially makes culture more uniform in multiple countries as ubiquitous media presence catapults artists to the top of the charts on every continent. But there is anecdotal evidence that small, regional music scenes can build themselves up by leveraging technology and global markets. This upward mobility can be achieved through the cross-cultural exposure that technology and globalization allows. One example of this is Kuduro music from Angola. Once a critical mass of popularity was reached, aided by the Internet, Kuduro groups began to tour in Europe which allowed them to accumulate capital which facilitated further development in their home, regional market. So, this is an example of technology and globalization fostering a niche (Diplo 2010).

II. History of the Recorded Music Industry and Review of Empirical Work

A closer inspection of the evolution of the music industry helps to make sense of the institutional changes that are occurring today. This is followed by a discussion of intellectual property (IP) rights in the recorded music industry and the associated appropriability regimes and appropriability mechanisms that are affected by the institutional changes. Finally, this chapter surveys the most relevant literature which has a similar research agenda. This situates the context for this study and sheds light on the uniqueness and importance of this research.

A. Evolution of the Recorded Music Industry

First, a brief explanation of contractual arrangements will help to explain the structure of the industry referred to throughout this research. The entity of central importance, creatively, is the artist or group of musicians referred to as a band. The band may write its own music and lyrics or may purchase music and lyrics from an external composer (Connolly and Krueger 2006). Other actors play a major role economically and legally. Managers or agents represent the band and perform other managerial services for which they take a portion of the band's revenues. Bands make contracts with concert promoters for live shows usually via the band's agent or manager. Concert promoters, also

called concert organizers, essentially employ the band when they are on tour, advertise the concert events, and obtain the venues. If and when an artist or band reaches a certain threshold of popularity, they sign a contract with a recording company to produce, market, and distribute CDs. Record companies may also promote concerts, though this is becoming less prevalent (Krueger 2005). Contractual arrangements between bands and concert promoters are varied, though they may be likened to book contracts in that there is an advance payment and, if sales exceed a certain level, the band receives royalties. A band contracts with a publisher if it composes its own music in order to copyright it. The publisher contracts with a performing rights organization (PRO), also referred to as a copyright collecting agency or copyright collective. A PRO acts as an intermediary facilitating the flow of royalties between copyright holders and licensees. For example, the PRO licenses music for radio stations, TV, restaurants, and malls, keeps track of the use of music, and collects and distributes royalties. Half of the royalties go to the publisher, the other half goes to the composer, and the manager may accrue a small portion of the composer's half. When a band is involved with a record company, it is often a long-term contract stipulating the advance, royalty threshold, and royalty rate. The bargaining power is in the hands of the record company. The advance to the band, for example, is often used to pay expenses which the band must shoulder like recording costs. These contractual arrangements and bargaining power have been gradually changing since the new century began (Krueger 2005).

By 2010, revenue streams from live music performance took precedence over revenue streams from recorded music. As the decentralization of the music industry middlemen and the democratization of music consumption have occurred, the intermediary agents in the old business models must compete to survive not horizontally but vertically in

the supply chain. The intermediaries between the consumer and artists, such as record labels, collecting societies, copyright collectives, and publishers, are all questionable in their relevancy. So, there will likely be winners and losers. External to the Internet, in the non-digital world, retailers pay royalties to the record label via publishers. Online, in the digital world, the retailer pays the record label directly and the record label pays the songwriter, publisher, or the Harry Fox Agency. In the former business models, cross-subsidization – where a high price charged to one group makes up for a low price charged to another – occurred because recorded music was still the lucrative revenue stream (Page 2007). Live music ticket sales helped to promote record sales. In the new business models, recorded music is given away for free and helps to promote concert ticket sales. The cross-subsidization has deteriorated and possibly reversed. Record labels may increasingly take over the job of concert promoter. This kind of arrangement is often referred to as a 360 degree contract. In exchange for promotion, marketing and other services the label takes a percentage of merchandise, concessions and other revenue streams which historically have gone to artists. The opposite arrangement is becoming more prevalent as well where the concert promoter takes over the job of record label, as seen in the arrangement between artists such as Madonna and Shakira and concert promoter Live Nation. In both cases, economically the record label has become obsolete in its original capacity.

Technological change has always impacted the music industry. In Peter Tschmuck's (2006) analysis, changing conditions of production have an effect on how different actors at different parts of the value added chain make money. But, especially in the last 100 years, the aesthetic styles which gain popularity in an era are also determined by technological change. This is because each institutional change which occurs engenders a change in the logic of music production, distribution, and reception. For example, with the rise of radio in the

1920s, the broadcasting industry gained prominence over the publishing and phonographic businesses. Along with it came the Jazz Revolution – the corresponding aesthetic style for the new conditions of production. The phonographic industry later gained importance again in the 1950s. Along with it came the Rock 'n' Roll Revolution. The industry has undergone dramatic shifts over the past decade because of advances in digital technology. This Digital Revolution is still unfolding and the aesthetic style or styles which take precedence will be enabled by this technology, particularly computer software.

The turn of this century brought digitization – the process of transforming goods and services from analog to digital form. Recorded music is no longer in a physical form or 'phonogram'. Most recorded music is now data, where reproduction and distribution costs are approaching zero. MP3 technology has allowed for time and space compression of digital music as well as cheap and near-perfect digital copies from CDs. Recorded music can be disseminated potentially infinitely through illegal downloading and illegal file sharing. The first instances were networks like MP3.com and Napster, both of which began in the late 1990s. Unfortunately, from an economic perspective, this 'piracy' has compromised the private good attributes of rivalness and excludability. Scarcity has given way to abundance (Anderson 2006). Recorded music now approximates a public good where most consumers are able to free-ride. This represents, by definition, a market failure.

Figure 3 demonstrates the decline of the entire recorded music industry since the turn of the century in dollar terms. The data in Figures 3, 4, and 5 comes from the Recording Industry Association of America (RIAA), which acts as a copyright collective, administers certifications for sales numbers, and sets physical quality standards for recordings. The numbers from the RIAA are based on manufacturers' year-end unit shipments multiplied by recommended retailer list prices, and expressed in millions of 2000

dollars. Net returns are accounted for, but the values do not represent sales per se. As illustrated in Figure 4, CD shipments have fallen and this is reflected in many brick-and-mortar retailers going out of business. Figure 4 demonstrates the dollar value of shipments of different formats of recorded music. The decline of record stores and floor space for CDs at big box media stores can be seen in Figure 5. Figure 5 illustrates the different marketing channels that the RIAA tracks. Digital downloads, in Figure 4, and Internet marketing channels, in Figure 5, have become more important over time, but are still floundering in dollar terms.

It is difficult for legitimate Internet business models to compete with free MP3s now available for file sharing on bit torrent. New forms of legal consumption of MP3s have not been very lucrative, and music downloads have often been utilized as a loss leader. For example, iTunes has served this purpose for selling iPods, and Rhapsody has served this purpose for Verizon's mobile phone service. Before digitization, copyright and the enforcement of IP rights constrained consumption, dissemination and production, but also creativity (Demers 2006). But from a broader perspective, digitization has brought about costs and benefits which are asymmetrically distributed throughout the music industry. According to one study, the new appropriability regime has hurt the revenues of the successful 'superstar' artists more than unknown or new artists (Blackburn 2004).

There is debate about what will happen to the structure of the industry and the distribution of revenues. Will the existing business models of incumbent firms/artists be supplanted by new business models of entrepreneur firms/artists? A survey of history reveals that independent – 'indie' – artists, indie labels, and indie companies have been an entrepreneurial force in the music industry. In the 1920s, the Jazz Revolution was enabled by the emergence of new, smaller record companies and radio stations. These relatively small

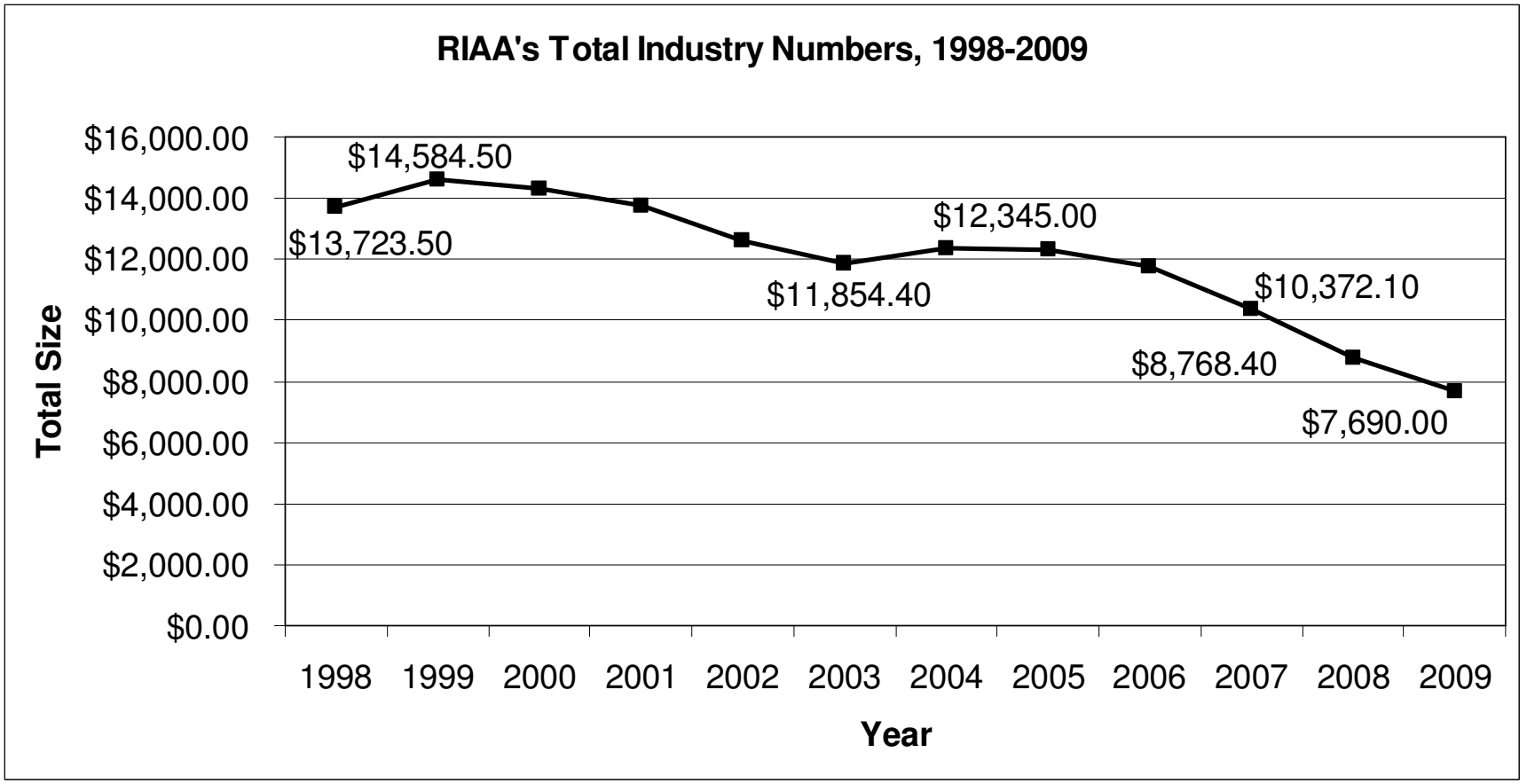

Figure 3: RIAA's Total Industry Numbers, 1998-2009.
Size of U.S. sound recording industry (millions of 2000 dollars). Source: RIAA.

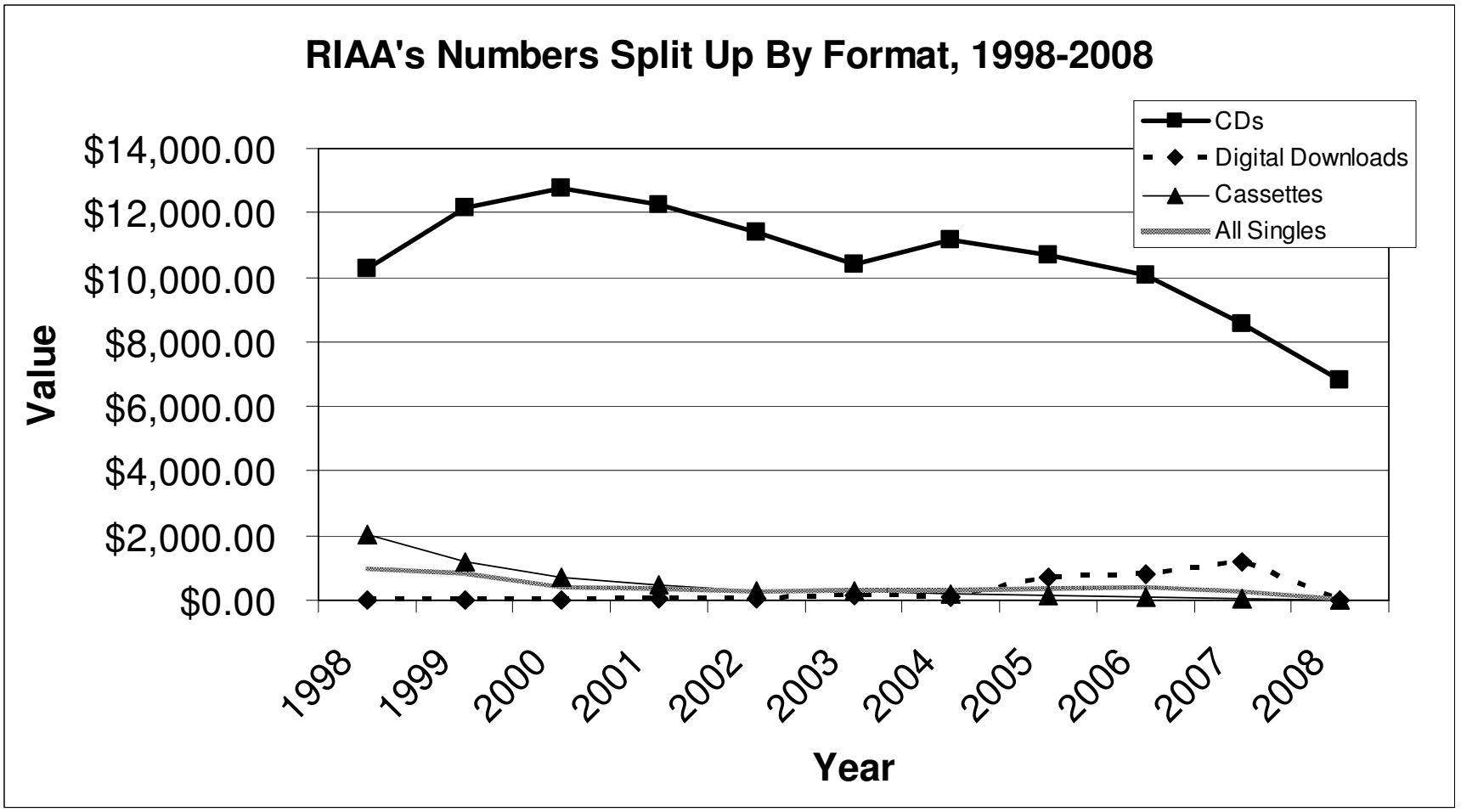

Figure 4: RIAA's Numbers Split up by Format, 1998-2008.
Comparison of full-length CDs, digital downloads, cassettes, and *all* singles (millions of 2000 dollars). Source: RIAA.

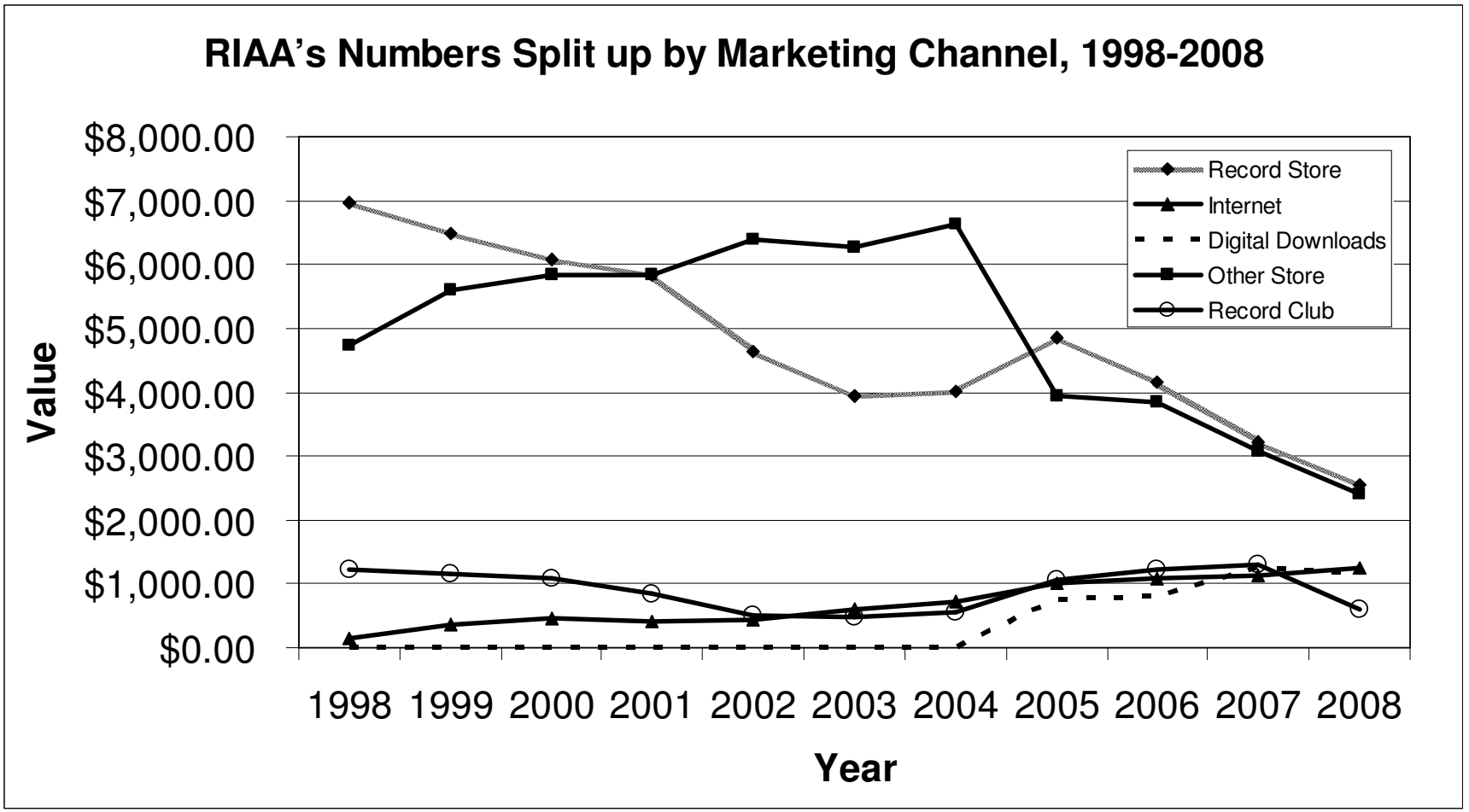

Figure 5: RIAA's Numbers Split up by Marketing Channel, 1998-2008.
Comparison of record stores, Internet, digital downloads, other stores, and record clubs (millions of 2000 dollars). Source: RIAA.

firms have often taken on the risks of developing new artists and finding new creative styles. This is because they have lower costs and can earn a quicker return on units sold than a large record company. Thus, they have a greater ability to 'roll with the punches' and adapt to changing conditions of production, distribution, and consumption. Indies can put more time investments into developing new acts, due to lower opportunity costs. Creatively, indies are more accurately tuned to their target audience's demand which facilitates entrepreneurship (Tschmuck 2006). Independent companies, as outsiders, have the ability to adapt to new technologies compared with incumbents who may be slow to transform their entrenched business models. Indies also circumvent the constraints that distribution monopolies and oligopolies place on creative expression and can therefore find new markets more easily. Leeds (2005) argues that, recently, entrepreneurial indie companies have gained market share from the incumbent majors due to technology changes.

Indeed, the thesis of creative destruction – from the earlier work of Joseph Schumpeter – explains that the entrepreneurs who cause creative destruction are characteristically small firms outside of the circular flow of existing production activities (Schumpeter 1911, cited in Audretsch 2008). So, the innovation which creates economic growth begins with the outsiders. This makes indies a very fitting agent of entrepreneurship in the music industry, consistent with Schumpeter's framework. What role will independents play now? This will depend upon how the structure of the music industry changes. A change in the distribution of revenues from incumbents to entrepreneurs would imply increased creative freedom and innovation. On the other hand, investment in new talent could be dampened due to the incumbent firms' heavy losses. Both effects could cause innovation as artists cooperate, defect, or shift to different modes of appropriating royalties.

Myopia hurt the incumbent firms because digital technology and the Internet made their business models obsolete. The major record companies did not adjust to the changed environment they were operating in. Digital technology, in the form of digital audio tapes (DATs), CDs, and MP3s, were all available by the early 1990s. But, by the time the technological possibilities all converged with high-speed connections, personal computers, the proper software, and the Internet, it was too late for the major labels. The major companies have exhibited a pattern over the last century of ignoring new technology, and then once it becomes a threat, fighting the new technology. This was true in the last century with both phonograms and radio broadcasting. Recently, MP3s and file sharing are the harbinger of the end of old business models. Ironically, this happened right at the apex of the incumbent firms' market success: teen pop, such as NSYNC, exemplified perfection in manufacturing blockbusters (Anderson 2006). At this time much capital was tied up in physical infrastructure to manufacture and distribute CDs, which now represents a sunk cost. But networks, bolstered by college campus Ethernets, further enabled the availability of free music at the turn of the century.

Recent changes in the recorded music industry have been devastating to old business models. There is potential for the industry to restructure in a decentralized way which fosters the relationship between the fan and the band. If artists and bands are free from constraining relationships with large companies with distribution monopolies and marketing bottlenecks such as terrestrial radio, then artists and bands can create their own rules. According to the French indie band Phoenix, there is 'no one between you and your fans' and there is more possibility for 'indies to make the pop charts' in the new paradigm (Phoenix 2010). But the broader context of the conditions of production in the music industry presents a new issue:

the absence of scarcity goes hand-in-hand with the absence of property rights. This diminishes the chances for remuneration and may adversely affect the incentives for artists.

B. Technology and the New Property Rights Regime

The year 2009 marked the one-hundredth anniversary of the Copyright Act of 1909 which provided the legal basis for the collection of mechanical and performance royalties by artists. In 1912, songwriters and publishers created the American Society of Composers, Authors, and Publishers (ASCAP) to license performance rights to broadcasting outlets and to distribute royalties to members. Performance royalties from radio play and live performances were paid to the copyright holders – songwriters and music publishing companies. In 1927, the Harry Fox Agency was created to license and collect mechanical royalties on phonogram sales for artists and labels. Mechanical royalties were paid to the recording artists, with substantial costs deducted. Preceding the digital and Internet revolution, the most lucrative revenue streams were from mechanical royalties on recorded music.

Digitization has allowed for circumvention of the IP rights regime, i.e., the copyright protection expressed in statutory and legal structures. These IP rights are government-granted monopolies which provide owners exclusivity in reproduction, performance, and derivative use. The combination of mechanisms that secure remuneration from consumers by strengthening a private good's attributes of excludability and rivalness is called an 'appropriability regime'. Excludability is the ability to prevent consumer access without payment. Rivalry means one consumer's use prevents another's use. For example, digital rights management (DRM) constrains consumer use of digital music which increases

rivalness. The new business models which utilize the available technology, like social network media, downloading, and streaming services to market recorded music can also be considered appropriability mechanisms. By simply not releasing specific albums or parts of albums on digital platforms, producers can deliberately create scarcity and excludability. The absence of The Beatles' music on iTunes is an example of a deliberately created scarcity, even though the music may be available in other digital formats. Even utilizing a band as a brand for differentiation or marketing merchandise is an appropriability mechanism.

However, the regime is tightened or loosened by the effectiveness of the appropriability mechanisms in place. For instance, technological possibilities often outrun the legal and statutory structure which protects IP rights. By raising the cost of enforcement of IP rights and lowering the cost of violating these rights, digitization has recently loosened the appropriability regime. This has decreased remuneration accruing to producers and content providers while increasing illegal consumer access to recorded music. Reactions to this, such as the Digital Millennium Copyright Act (DMCA), The Performance Rights Act, and Recording Industry Association of America (RIAA) litigation, are aimed at tightening the appropriability regime. Regardless, the loosening effect allows consumers to pay a much lower price than under the previous regime, which involved buying physical analog records.

Digitization brings about costs and benefits which are asymmetrically distributed throughout the music industry. Digitization and the Internet have hurt the revenues of the successful superstar artists more than unknown or new artists (e.g. Blackburn, 2004). This is the result of rampant illegal file sharing and downloading which substitutes for higher-priced and higher-margin CDs. Legal downloads have seen increasing importance, but at lower margins. However, it is difficult for any business model to compete with free MP3s. As mentioned, this has an asymmetric effect. Superstars incur heavy costs from the

dissemination of their recordings for free. On the other hand, this is where unknown or new artists benefit. Free marketing with free MP3 dissemination helps them to overcome the hurdle of getting known and expanding their fan base. But once an artist builds up a fan base beyond a certain threshold or network size, it is in their interest to protect their recorded music. Upward mobility and commercialization leads to increased incentives for copyright protection. Some popular artists put much energy into copyright protection and enforcement, such as Metallica. Others posture themselves as disinterested in commercial gain, such as Radiohead (Bourdieu 1993). But the asymmetric costs and benefits, from an economic standpoint, remain.

Mechanical and performance royalty payments are differentially affected. The focus now is on live performances because they are the most lucrative income source. Contracts between artists and record companies and between artists and concert promoters are looking much different than they were less than a decade ago. The increased utilization of '360 degree' contracts means that record companies are getting larger shares of revenue streams which formerly went to artists, such as merchandise sales (Leeds 2007). But this also means that, in many cases, record companies are losing viability as phonogram sales and mechanical royalties slide relative to performance royalties.

Retailers have been scrambling to devise new strategies which cater to consumer preferences for Internet downloading rather than traditional record store purchases (Demers 2006). This is seen in the changing popularity of marketing channels for recorded music, as seen in figure 5, and a significant increase in supply and variety. Online retail has several advantages over brick-and-mortar retail. This is observed in other industries, such as books, by comparing the business models of Barnes and Nobel to Amazon.com (e.g. Anderson 2006; and Brynjolfsson, Hu, and Smith 2003). Online retail is not constrained by geography,

and sellers can reach more customers over the Internet. Physical shelf space is also not a constraint for online retailers, so selection can be much broader due to decreased inventory costs. Inventory is nonrivalrous on the Web, and often there is too much for consumers to navigate through. This problem is remedied by information tools and better filters. Algorithm-fueled recommendations, by decreasing transactions costs, help customers find titles and create positive sales externalities (Anderson 2006). The lack of distribution bottlenecks allows Rhapsody, a subscription-based MP3 streaming service and retailer, to provide customers with back catalog, older albums, live tracks, B-sides, remixes, covers, foreign bands, and obscure bands on obscure labels.

In sum, technological change has transformed institutions in the recorded music industry. The technology change has occurred with the help of personal computers, software, and the Internet. In a pattern relived throughout the past hundred years in this industry, technology has changed faster than the legal structure. This has asymmetric costs and benefits for different agents. The superstar and long tail effects are likely reflections of this asymmetry. Is the pie – total industry revenues – growing? Is it being sliced differently? This analysis is of great importance for the future. If there is a long tail effect, increased diversity and a leveling of the playing field is occurring. If there continues to be a superstars effect, increased homogenization on the few at the top is occurring. Finally, if innovation is still coming from the outsiders – the indies – then new creativity will certainly earn money differently in the new regime.

C. Literature Review of Empirical Research

The hypotheses outlined in the first chapter have been empirically examined by others studying various industries. Anderson's long tail thesis is applicable to multiple markets. However, prior research has not looked at all three hypotheses together. Additionally, other research has not investigated such a recent time frame as the one utilized here. Nielsen Soundscan data is not often used due to its proprietary nature. My research is timely because it will draw from the methodologies used below to examine the long tail and superstars effects in the recorded music industry while there are lively debates occurring in industry groups and academic circles.

Chevalier and Goolsbee (2003) acquire sales and sales rank data for online books sellers. They assume a Pareto distribution, which is more generally an exponential function such as:

$$\mathbf{Pr(s>S)=(k/S)^{\theta}},$$

and regress log(rank – 1) on log(sales):

$$\mathbf{(Rank-1)/(Total\ \#\ Titles)=(k/S)^{\theta}},$$

and alternatively expressed:

$$\mathbf{\ln(Rank-1)=c-\theta\cdot\ln(Sales)}$$

Assuming a power law distribution, this is a regression of sales rank on sales. This allows estimation of Pareto slope parameters. I will follow this method to help determine the distributional shape of my data. These authors do not test the hypotheses I will use. They use their parameter estimates to calculate degrees of market power and inflation.

Elberse and Oberholzer-Gee (2008) test both the superstars and the long tail hypothesis. They collect data for movies from *Nielsen VideoScan* for sales and sales rank to analyze the allocation of revenues across time. They find a significant change in sales distribution over their sample period which shows a 'flattening' of the tail, which means a larger number of low-volume titles sold, consistent with the Long Tail hypothesis. Simultaneously, they discover that 'superstar' movies did worse each year over the sample. But, contrary to the Long Tail hypothesis, the larger number of low-volume titles were not making up for the diminishing sales of high-volume hits. The variety of supply, or number of products, had definitely increased, but there was also a two-fold increase in titles with zero sales. Additionally, there were ever fewer titles at the 'Head' (or most-popular) end of the Long Tail distribution which meant increased concentration of 'Superstar' sales on fewer titles. This is consistent with the Superstars hypothesis to the extent that there is increased *relative* sales in the Head of the distribution – slicing the pie differently – but there is also a decrease in the absolute size of sales – decreasing the size of the pie.

The authors use nonparametric tests to analyze the distribution of revenues across titles over the time period of their sample. They calculate location, scale, skewness, kurtosis and create inter-quartile measures to pick up how the 'Head' or 'Tail' shifted in relation to the median. This allows division of the distribution into different groupings with equal frequencies such as quartile, quintile, etc. They find evidence that the distribution becomes

more heterogeneous, but also more skewed and asymmetrical: a sharper peak *and* a longer tail. This supports the hybrid hypothesis.

Many factors can explain the above changes, and non-parametric results cannot control for product characteristics or changes in buyer preferences, so they also estimate quantile regression models to test the effect of other business environment changes on the sales distributions. They use this model because the mean is not of interest in this type of analysis, but distributional changes are. Essentially, this is a linear regression with quantile as the conditional dependent variable:

$$Q_\theta(y \mid x)=x'\beta(\theta)$$

Their independent variables are a set of yearly indicators. The coefficients on these decrease for a quantile in the 'Tail' of the distribution if there is a Long Tail effect. They find the distribution of sales has shifted down in general, but there was also a shift in the mass towards niche products. Besides the Long Tail thesis, explanations like a shorter product life cycles could account for declining sales per title in the Head of the distribution. To test if the number of titles reaching certain sales thresholds has changed over time, they also estimate a negative binomial regression model. This method controls for the total number of titles available in a given year, which is a potential supply effect explanation. They find a flattening and a lengthening of the Tail. Quantile regressions show falling sales for superstars, and the negative binomial regressions show that this superstar quantile had half as many titles as compared to the beginning of the period, consistent with a superstar effect. Together, these results support the hybrid hypothesis.

Similarly, Page and Garland (2009) find support for the hybrid hypothesis in legal and illegal digital music markets – a combination of the long tail and superstars hypotheses. For legal digital music, they find a very long tail, but with 75% of total inventory not selling at all. There is simultaneously a horizontal compacting of the head of the distribution going to a shrinking number of hits. Within the digital music black market – peer to peer (P2P) usage data – they find a similar distribution, consistent with the hybrid hypothesis. P2P has a longer tail than the legitimate market. However, they only analyze their observed distribution in a Lorenz curve framework. This does not consider absolute changes. The pie is sliced differently, and the size of the pie has shrunk for legitimate markets. But they are unable to observe the 'size of the market' for P2P.

Brynjolfsson, Hu, and Smith (2003) also use data on sales and sales rank from an online bookseller, Amazon.com, fit to a log-linear distribution:

$$\mathbf{\log(Quantity) = \beta_1 + \beta_2 \cdot \log(Rank) + \varepsilon}$$

A log-linear curve is a Pareto curve. A Pareto distribution follows a power law. They regress log(Quantity) on a constant term and log(Rank), effectively testing the long tail and superstars hypotheses. This allows estimation of coefficients, 'Pareto slope parameters', which are found to be negative and significant. The estimates are used to calculate the proportion of unit sales which reside above particular sales ranks relative to the total number of available book titles. They find support for the long tail hypotheses. From their parameter estimates, they calculate consumer surplus gains from online marketing channels.

Brynjolfsson, Hu, and Simester (2007) use Gini Coefficients and Pareto Curves to measure sales concentration of a retailer with Internet and non-Internet catalog channels.

Again, this is effectively testing the long tail and superstars hypothesis. They find the Lorenz Curve for the Internet channel is less skewed towards the most popular products than for the non-Internet channel. The Gini Coefficient for the Internet channel is lower than the catalog channel. This information gives estimates of percentages of sales generated by different percentages of products. But statistical significance cannot be obtained from these estimates. So, they also find Pareto slope parameters by regressing sales onto sales rank. Connolly and Krueger (2006) utilize the same model for concert sales. Brynjolfsson, Hu, and Simester (2007) estimate a log-linear relationship:

$$\ln(\text{Sales}) = \beta_1 + \beta_2 \cdot \ln(\text{Sales Rank}) + \varepsilon$$

They find the difference between the coefficients of the two marketing channels is significant. They then re-calculate the Lorenz Curve and Gini Coefficient, controlling for selection bias. The Internet channel Lorenz Curve lies above that of the catalog channel, so the Internet channel's Gini coefficient is smaller. This implies the sales distribution is less skewed for the Internet channel relative to the catalog channel. They then repeat the estimation of the sales and sales rank model above with similar results, controlling for differences in consumers. The β_1 coefficient is significantly less negative for the Internet than for the catalog channel, which implies the sales distribution has a longer tail in the Internet channel. This supports the long tail thesis. The authors also utilize data from individual consumers to determine that when consumers have lower search costs this leads to a sales distribution less skewed toward hits – less concentrated in the 'Head' of the distribution.

Original observations of stardom came from Irving Fisher and Alfred Marshall (Fisher 1925, Marshall 1947, cited in Hamlen 1994). As average income increases, the demand for high-quality goods and services rises faster than the demand for low-quality goods and services. Improved technology decreases production costs and lowers output prices, which allows higher-quality products to consolidate market share. In Marshall's time, this did not apply to music to the degree it did for Rosen (1981). For Rosen, skewness in earnings was the result of differences in talent and imperfect substitutability. Subsequent research attempted to test how this might come about. McDonald (1988) extended Rosen's analysis into a dynamic information accumulation process of popularity.

There are differing ways to define popularity or success. Popularity within a genre and various chart measurements are used. Chung and Cox (1994) use number of Gold Records. Because Rosen's specific formulation of the hypothesis relies on talent differences, talent is defined in different ways. Hamlen (1991), for example, uses an objective measure of voice quality to proxy for talent. Krueger 2005 uses the volume of articles written about particular artists.

Factors other than talent could lead to a superstar effect. The social nature of music, for instance, may lead to a 'snowball effect', or increasing returns to success, where a highly skewed sales distribution could emerge regardless of talent. Artists with equal talent could experience vastly different sales levels due to unique properties of consumer demand for recorded music. Adler (1985) argues that consumers minimize search costs. Music consumption requires knowledge and consumers can only accumulate a limited amount through discussion. They therefore specialize in certain artists or genres to minimize costs. Similarly, Crain and Tollison (2002) argue that consumers economize on accumulating the knowledge – 'consumption capital' – that would allow them to enjoy the consumption of

music, which leads to a superstar effect. Hamlen (1991) contradicted Rosen's analysis by showing that increases in his 'objective' measurement of talent – harmonic content – show sales increasing at a decreasing rate. Consumers could, theoretically, recognize talent – singing quality – but increases in talent lead to less than proportional increases in record sales. Decreasing returns to ability is the opposite of what Rosen suggested. Chung and Cox (1994) show how a superstar distribution, through specific kinds of random probability mechanisms, could arise regardless of talent. However, Giles (2006) uses the same method with different data and finds slightly different distributions. This lends some credence to Rosen's original idea. Different methods are used by Spierdijk and Voorneveld (2007) to test different probability mechanisms. Though different stochastic processes are tested out, all of these studies imply that luck or random factors could lead to large inequalities in popularity.

D. Summary

These examinations of the superstar phenomenon have some problems. Any 'objective' measurement of quality, talent, or ability is problematic because any conclusions are too sensitive to the scale chosen. Even though Rosen relies on differences in talent for his explanation, it is futile to define it. For example, independent, 'indie', fans have different definitions of what 'indie talent' is and what legitimate locations to see indie concerts are. Progenitors of different styles of music, genres, and sub-genres utilize numerous different cultural symbols and have numerous sources for legitimation and consecration which are unlikely to be agreed upon. There are combinations of factors which lead to the distribution profiles which are observed. Certainly, random factors do play an important role in which artists becomes popular. Consumers and the social and cultural nature of music

consumption are important; but the talent, ability, and productivity of the artists/bands play an important role as well. Investment in marketing strategies and promotion are a huge factor in who gets exposure, radio airplay, and visibility to mass audiences. It is a confluence of these effects which produce superstardom.

It is unnecessary to attempt to define an objective measure of talent. For example, talent is multidimensional (e.g. Crain and Tollison. 2002) and inter-genre comparisons of talent are problematic. It is also problematic to model consumer behavior in relation to cultural goods. First of all, the data used in this study does not observe consumer attributes. Beyond this, consumption and production of cultural goods involves competition for symbolic distinctions and symbolic profits (Bourdieu 1984, 1993). Because it is symbolic, it is difficult to define commensurate scales of comparison. This cannot be modeled by utility maximization or cost minimization as many researchers have done.

A shift in supply has been brought about by technological and institutional change. This is altering the structure of the industry and eroding property rights. Consequently, for this research the sales distribution is examined across the sample period to test if sales concentration has increased or decreased.

Therefore, this study incorporates the methodologies of the studies outlined above which simply analyze outcomes. The presence of changes in the sales distribution profile is evidence enough to accept or reject the superstar, long tail, and hybrid hypotheses as they are defined here. This does not require arbitrary and questionable assumptions about objective quality, or consumer behavior in cultural markets. Though the data utilized for this research is the most accurate unit sales information in the recorded music industry, the data shortcomings outlined in the next chapter also necessitate only analyzing distributions of the sales outcomes.

III. Data Collection and Data Description

This chapter summarizes the data collection process and describes the resulting data set. The first section analyzes the procedure of creating the data set and the problems encountered in that process. There is also a description of the specific formats that are tracked. In the second section, graphs and descriptive statistics are used to shed light on the basic patterns that can be observed in the data set.

A. Nielsen Soundscan and Data Collection Procedure

The data for this study comes from Nielsen Soundscan. Nielsen Soundscan collects point-of-sale transaction data on unit sales from over 14,000 retailers across the U.S. It comes from retail, mass merchant and non-traditional distribution channels (on-line stores, venues, mail order, and digital services). Big retail chains, such as Wal-Mart, Tower Records, Virgin Megastores, and Sam Goody – now For Your Entertainment (F.Y.E.) – report to Soundcan. The sales information collected by Nielsen is used by most music industry operators and has been used to create the Billboard music charts since 1991. They provide academic access packages as well.

The necessary information for my research from the Nielsen system, pulled from 'Chart and Sales History' reports, are only available starting in 2004, due to a Nielsen system

upgrade that July. But, the available time frame begins in week 1 of 2004 and goes through week 52 of 2009. My focus is on a random sample of artists available during this time frame. Because of truncation, the random sample covers 2004 through 2008. This is due to data problems which did not allow time to complete the retrieval of all 2009 sales information for previously-gathered (prior to 2009) and newly-released 2009 artists and titles as this year-end information became available. Consequently, to make the sample size as large as possible, all 2009 sales were removed and we focus only on five years, 2004 through 2008.

Artists were chosen randomly by a random letter generator and random number generator. The sample, then, does exclude artists and bands whose name begins with numbers. An English letter and subsequent page number were both chosen from a discrete uniform distribution and the associated page corresponding to that letter-number combination was located. As time went on, the number of pages (from which to draw a random number) for any letter of the alphabet changed as new titles were released by artist names beginning with that letter further up in the alphabetical list. So, this required the number to be updated depending on the number of titles available. A search query in Nielsen Soundscan's title search for a given letter brings up a specific number of pages (for certain letters, thousands) of results in a search query. Specifically, my 'title lookup' selection criterion was search for: 'letter', search by: 'artist', options: 'match from beginning', and format: 'album/single/video'. Nielsen Soundscan lists artists alphabetically and each webpage presents 100 'titles' for those artists. Individual 'titles' correspond to the names of albums, singles, and videos. Sales histories for 'digital tracks', as opposed to what I refer to as 'unbundled tracks', are organized differently in the system compared to albums/singles/videos because they have their own products codes independent of the album or single. These were therefore excluded for the sake of brevity as our access to

Nielsen was limited. Digital tracks, however, would be interesting to examine in future research projects.

More specifically, after a random letter was chosen a random page number was chosen, depending on how many pages of 100 titles showed up on that date. Nielsen's search results list all titles belonging to alphabetically-listed artists ranging anywhere from one to 100 artists per page. This depends on how many titles were currently available for sale by those particular artists. My sample includes all titles available for all artists listed on a particular randomly-chosen page. So, if a particular artist's list of titles spilled over onto another page or pages, all of those titles were also included.

Individual 'artists' are defined by Nielsen Soundscan. An 'artist' could be an artist's stage name, an artist's real name, a band name, an orchestra conductor's name, a compilation of different conductors together, a movie soundtrack, a record label sampler, or other collaborations such as a movie or documentary for which copyright holders have an interest in tracking. There are many observed errors presumably committed in the process of retailers reporting information to Soundscan. I omit some artists who should not belong in the sample and correct misspellings. Some artists release titles with a band as well as solo. If the solo artist's name or stage name was listed on the page randomly chosen, this became part of the sample. But any other artist-band collaboration the artists might be involved with was omitted if it was not listed on the page. The reason for this is that an artist and band collaboration includes other creative inputs from other artists. It is the combination of inputs, i.e. defined by Nielsen's listed 'artist', which were randomly chosen and so belong in the random sample. So in general, an individual artist could be an individual artist's name or stage name with or without any collaboration by another artist, conductor, band, or bands. What was listed as the artist on the chosen page is what was included and nothing more. For

example, if an artist showed up on the randomly chosen page, but a release recorded by this artist and her band showed up on another page, only the solo artist is included. However, in many cases, the retailer reporting sales to Soundscan misspelled an artist's name, stage name, or band. The misspellings provide another layer of randomness to the sample. This is because a misspelled name would be listed on a randomly chosen page when the artist would not have been included in the sample if not for the misspelled name. To amend misspelling issues, I searched for the correct artist name and all of the common incorrect spellings of that name. At Nielsen's recommendation, I included unit sales information for both correctly and incorrectly spelled artist names and bands and listed them under the proper nomenclature, if it could be tracked down. For example, the stated proper listing for an artist's first and last name was LAST*FIRST. However, I encountered:

J DAWG
J-DAWG

CLARK*LAWRENCE J.
J CLARK*LAWRENCE

WU-TANG KILLA BEES
WU-TANG KILLA BEEZ

JOHNNY BLAS
BLAS*JOHNNY

WILLIAMS,DON
DON WILLIAMS
WILLIAMS*DON

Even if the name was listed incorrectly, and as long as the proper nomenclature was found, all artists were included with all different listings, correct or incorrect. For example, entries for 'artist' Johnny Cash included:

JOHNNY CASH
CASH JOHNNY
CASH JOHNNY (FR)(W/DVD)
CASH*JOHNNY
JOHNNY*CASH

In this case, the correct entry would be CASH*JOHNNY. However, these would be listed on different pages because search results were listed in alphabetical order.

I used the Chart and Sales History title reports because they displayed year-to-date (YTD) and release-to-date (RTD) unit sales broken into each sales format for each week. The broad format characterizations are albums, singles, or videos. These are then divided into YTD total units and YTD digital units. RTD total units information for album releases also includes more specific formats:

1) **LPs**. 'Long-playing record' albums are 33⅓ revolution per minute (rpm) analog phonograph records, usually having a diameter of 10 or 12 inches. Since 1948, this was the chief format for phonograms until the advent of the compact disc (CD). The phonograms can be made of vinyl (polyvinyl chloride) or not. 'Long playing' refers to the relative time capacity.
2) **Cassette tapes**. Also called audio cassettes, these contain an analog audio signal written on plastic magnetic tape spun around two spools.
3) **CDs**. A compact disc stores digital audio data, but is still a phonogram, in physical form.
4) **Digital video discs** (DVD). These contain a digital signal. DVDs are often bundled together with CDs. This form of bundling is not reported by Nielsen.

5) **Digital albums**. This is an album of individual cuts or tracks bundled and sold together via an online retail platform such as iTunes, Amazon, Rhapsody, etc. Sometimes individual tracks can be sold and downloaded separately or unbundled. These are categorized differently. 'Digital tracks' have their own serial number and are categorized differently than digital albums and unbundled tracks, but are not included in this sample.

Singles are shorter than albums with one or more tracks. Songs on a single can also come out on an album. Singles often contain a song or songs which are the most popular song or songs from an album and therefore serve a promotional purpose. RTD total units for single releases specifically includes:

1) **CDs**.
2) **12-inch singles** (12"). The 12-inch single analog-signal phonograph record is made from acetate and emerged from the disco era, circa 1970. These were originally smaller (10") but are usually now the same diameter as an LP, but with rpm of 45 and other characteristics which improve sound quality, particularly the base frequencies.
3) **Cassette tapes**.
4) **Digital singles**. This is a single of individual cuts or tracks bundled and sold together via an online retail platform such as iTunes, Amazon, Rhapsody, etc.
5) **Maxi CDs**. This is also referred to as a 'maxi single', and sometimes abbreviated 'MCD'. It is a music single release with more than two tracks in CD form. The origin is the same as the 12-inch single but at first it included more tracks.

RTD total units for videos specifically includes:

1) **UMDs**. A 'Universal Media Disc' is an optical disc medium made for the Sony PlayStation Portable which is a handheld video game system.
2) **DVDs**.
3) **Video tapes**. These contain an analog video signal written on plastic magnetic tape spun around one or two spools.

YTD and RTD unbundled tracks are also tallied for both albums and singles. Consumers often have the option from many online MP3 retailers to purchase one or more individual tracks unbundled from the 'single' unit or 'album' unit.

The first pages that were used included information from week 52 of 2004, 2005, 2006, 2007, and 2008 for each individual title. Any artist names (described above) which in my best judgment were misspelled or could be commonly misspelled were individually searched in the Nielsen system to find as many misspellings as possible. All of the unit sales information for alternative spellings of those artists was included as well as newly discovered misspellings. Chronologically, week 52 2009 information was not available until the end of our Nielsen access window. The previous week's data is available the following Wednesday. The first time around, I kept a record of the starting point, all of the artists and titles saved, and the ending point of the random page's location in the alphabetical listing of artists, so that I could find it again. This is because I anticipated the position of the artists and titles on the page I originally visited would change over time as new 2009 titles were released further up the alphabetical list as time went on. After downloading the latest information, I had to check once again for misspellings for the new artists which were picked up. Bounce Software LLC put the final html files together into excel form. The final excel file data needed further revision due to various issues. Many catalog numbers were duplicated for different releases of a title, so there was no criterion for a unique title, and any fixes were done manually.

Additionally, many titles had changing characteristics over time. For example, record label names sometimes changed from year to year due to mergers and other ownership changes. Artist and title names were sometimes spelled differently from year to year. Listed release dates also changed as time went on. This is likely due to reporting errors in Nielsen's system. Again, any fixes related to these issues were performed manually.

All together, this study utilizes a random sample of 7,010 albums, singles, and videos released by 2,051 different artists and 1,836 different labels to study the distribution of unit sales in the U.S. music industry from 2004 to 2008. The detailed level of the formats included here is important because of the technology changes that have been occurring over the last few years and shifting consumer preferences.

When a yearly sales figure of zero is observed for a given artist's title, it is not known whether this title is available for sale or not. All that is known is that at some point in the past, the title was registered with Nielsen Soundscan. A band, for example, could have broken up years ago. This is important because the number of titles and artists with yearly sales of zero doubles from 2004 to 2008. So certain analyses below are done for two groups. The first group is non-conditional sales, which includes the zeros. The second group is conditional sales, which drops the zeros. So 'conditional sales' for the second group is conditional on sales being greater than zero, or nonzero.

There is likely some degree of sampling bias due to the sampling methodology. Given the constraints present in the Nielsen Soundscan system, no sampling methodology could have made artist or title selection random with a uniform probability. Characteristics of the population were not known, or withheld by Nielsen. Using RIAA figures, the best estimation of the percentage of total industry unit sales that are present in this sample is 0.18%. The total number of titles and artists that Nielsen tracks was unknown. The large

volume of data in the system and the restrictions on search queries compounded this. Consequently, artists with larger numbers of titles are over-represented because they are present on more pages and more likely to show up on any given page of 100 titles. Artists with smaller numbers of titles are then under-represented. This suggests that the sample may be biased towards superstars because the number of titles is correlated with higher unit sales. Even if more information was known about the population from Nielsen and unique artists could be pulled, the spelling problems outlined above may still have created bias and made a random sample difficult to achieve. Artists who were spelled incorrectly are over-represented because their titles span more pages in the system and are more likely to be picked up. Additionally, the sample is biased towards artists who have been around longer because they have released more titles and product life cycles for each individual title have more time to work themselves out. Because superstars have longer product life cycles, the sample is once again biased towards them.

B. Description of the Data

Figure 6 displays all unit sales, bundled digital sales, and unbundled track sales from 2004 to 2008 for the sample used for this study. Non-digital sales are also shown, which is all sales minus digital sales. A detailed description of each format discussed in this research is found in Table 1 below.

Table 1: Summary of Recorded Music Formats.

Album Releases:	Single Releases:	Video Releases:
CDs or compact discs store digital audio data, but is still a phonogram, in physical form.	**CDs** (explained in left column)	**UMDs** or 'Universal Media Discs' are optical disc mediums made for the Sony PlayStation Portable which is a handheld video game system.
LPs or 'Long-playing record' albums are 33⅓ revolution per minute (rpm) analog phonograph records, usually having a diameter of 10 or 12 inches.	**12 Inch Singles** (12") are analog-signal phonograph records made from acetate which emerged from the disco era, circa 1970.	**DVDs** (explained in left column)
Cassette Tapes are also called audio cassettes, these contain an analog audio signal written on plastic magnetic tape spun around two spools.	**Cassette Tapes** (explained in left column)	**Video Tapes** contain an analog video signal written on plastic magnetic tape spun around one or two spools.
DVDs or digital video discs contain a digital signal. DVDs are often bundled together with CDs. This form of bundling is not reported by Nielsen.	**Maxi CDs** or 'maxi singles' or 'MCD' are music single releases with more than two tracks in CD form.	
Digital Albums are albums of individual cuts or tracks bundled and sold together via an online retail platform such as iTunes, Amazon, Rhapsody, etc. Sometimes individual tracks can be sold and downloaded separately or unbundled. These are categorized differently.	**Digital Singles** are singles of individual cuts or tracks bundled and sold together the same way as a digital album. Sometimes individual tracks can be sold separately.	

All unit sales rise from about 5 million in 2004 to peak at 7 million in 2006. From there, all unit sales decline to about 3 million – its lowest value – by 2008. The rise in digital unit sales increases consistently across this time frame from 43,000 in 2004 to 380,000 in 2008. Comparatively, digital sales are still dwarfed by CD sales. Digital sales are only 3% of CD

sales in 2004 and 14% of CD sales in 2008. So, digital albums do not compensate for the sharp decline in physical CD albums. But at the same time, unbundled track sales rise drastically from 3 million to 43 million. This exemplifies the increased consumer preference for individual songs over the album as a unit. This is facilitated by the record label and retailer strategy of mixed bundling (Elberse 2009). Figures 7 and 8 confirm the above trends with sales per artist and sales per title. The same patterns are observed when sales are divided by the number of artists and number of titles, respectively. However, average unbundled track sales do decline for both artists and titles in 2008. This decline is a symptom of the crowding of the market due to lower barriers to entry catching up to sales numbers. The supply of artists and bands has increased and so has competition.

The release dates for the titles in the sample are described in Table 2 for all albums, singles, and videos. The average release year is 2002. The median release year is 2004 which is the first year of the sample. The earliest release year is 1957 and the standard deviation is 5.5.

Table 2: Descriptive Statistics for Release Dates and Release Years.

Mean, median, standard deviation (in years), minimum and maximum.

	Mean	Median	SD	Min	Max
	37458.6368	38053	5.52836264	20821	39812
Release Date	7/21/2002	3/7/2004	5.52836264	1/1/1957	12/30/2008
Release Year	2002	2004		1957	2008

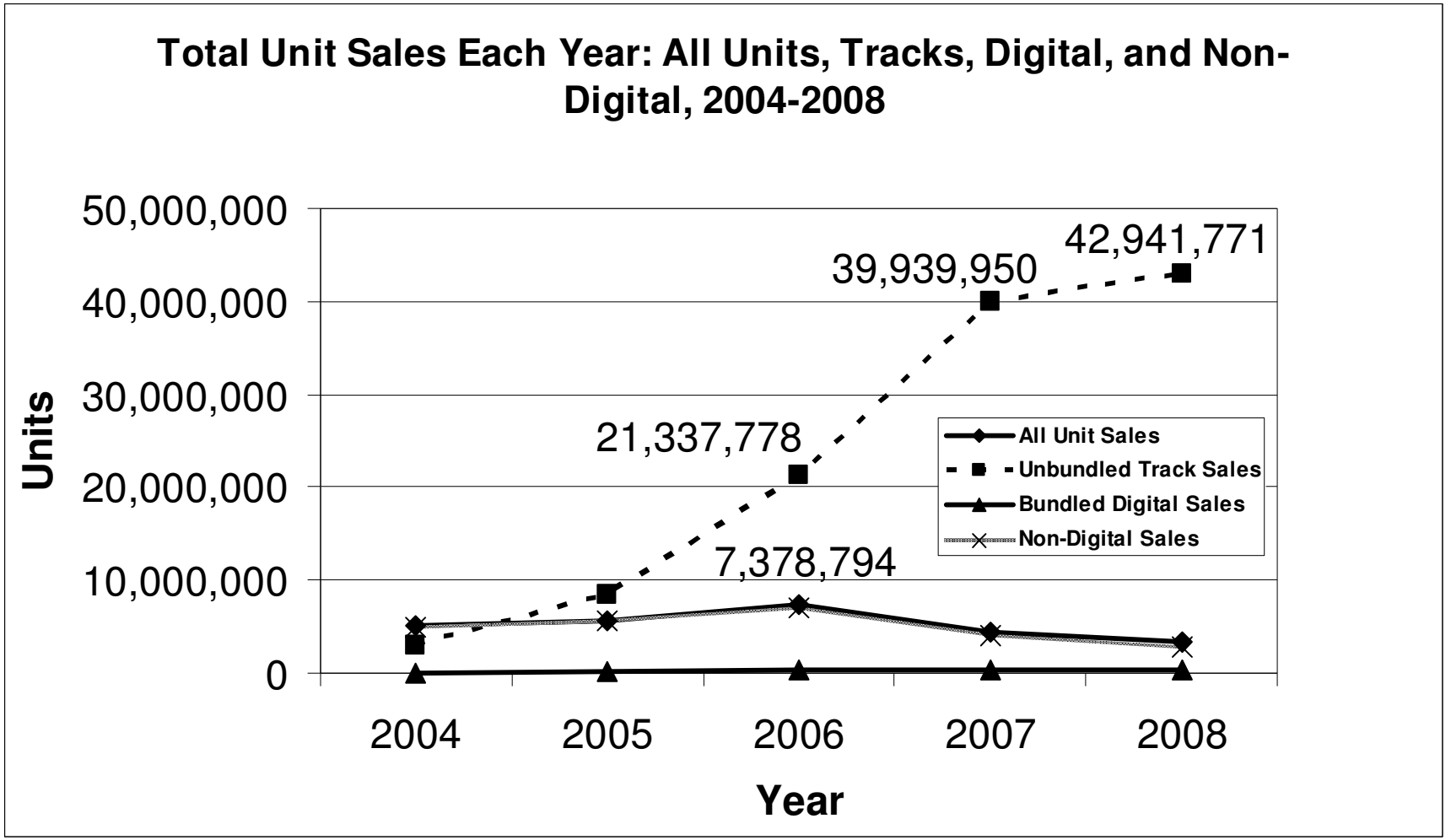

Figure 6: Total Unit Sales Each Year: All Units, Tracks, Digital, and Non-Digital, 2004-2008.
For each year of the sample period. Included are all unit sales, unbundled digital track sales, bundled digital units sales (i.e. digital albums or singles), and non-digital sales (all unit sales less bundled digital sales).

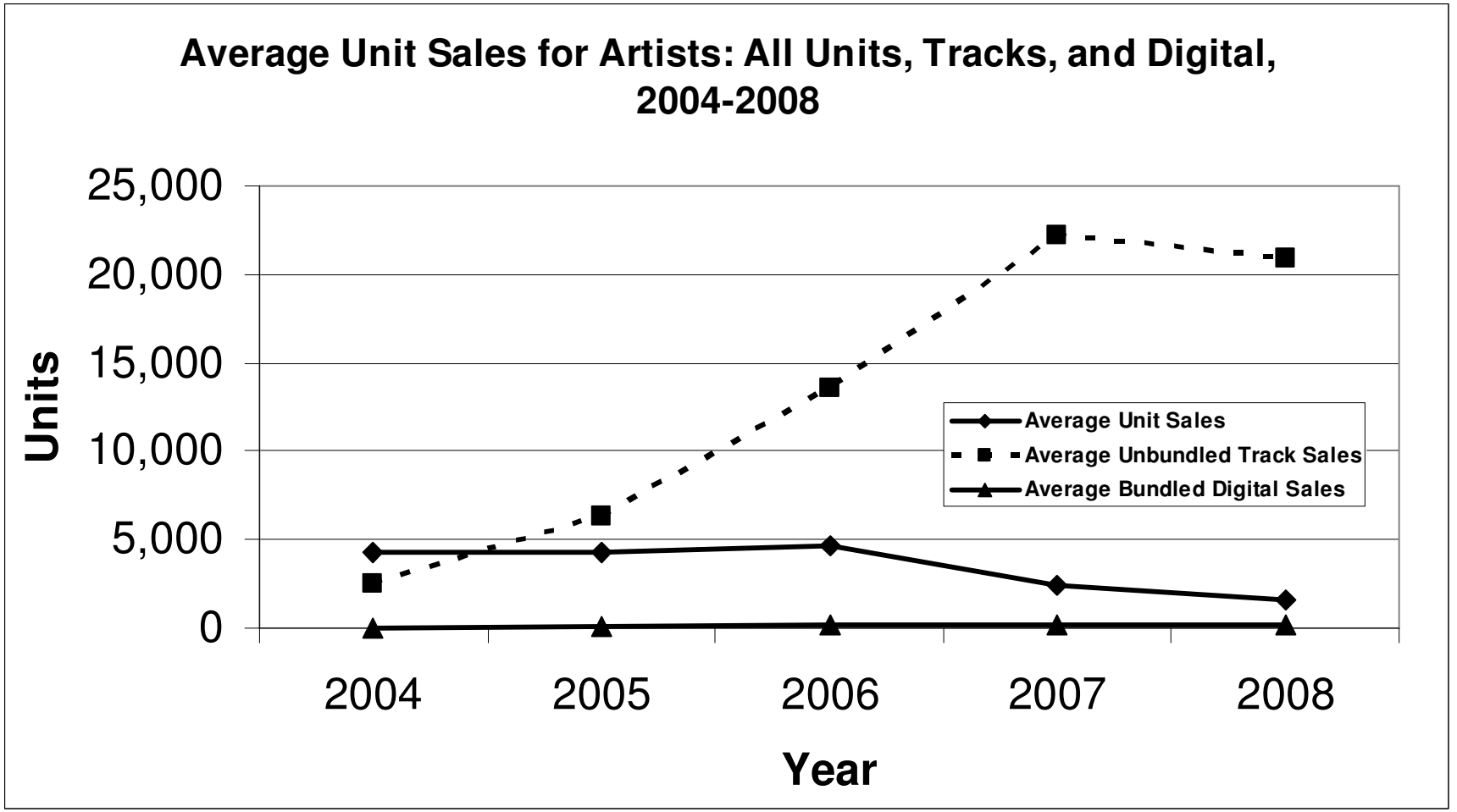

Figure 7: Average Unit Sales for Artists: All Units, Tracks, and Digital, 2004-2008.
For each year of the sample period. Included are unit sales, unbundled digital track sales, and bundled digital unit sales.

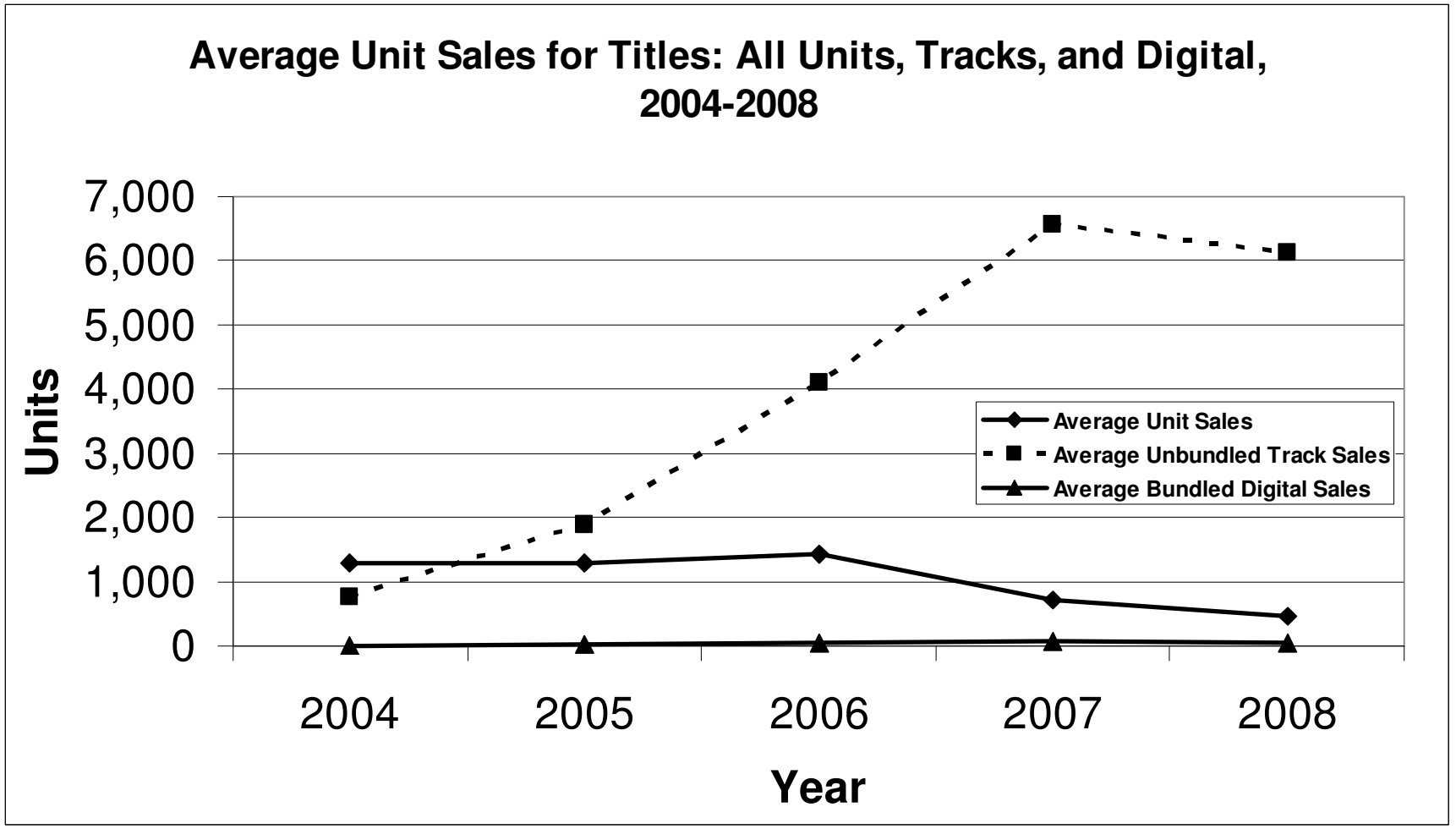

Figure 8: Average Unit Sales for Titles: All Units, Tracks, and Digital, 2004-2008.
For each year of the sample period. Included are unit sales, unbundled digital track sales, and bundled digital unit sales.

Table 3 presents some descriptive statistics for the different formats being analyzed. The general categorizations are albums, singles, and videos. There are 7,010 titles total, of which 6,349 are albums, 310 are singles, and 351 are videos. These are further broken down into more specific formats. The absolute numbers in the first column indicate how many titles sold greater than zero units of the specific format listed. The percentages listed indicate the fraction of the total. It is the same for albums, singles and videos. The list of specific formats is not mutually exclusive, i.e. does not add up to 100%. The percentages listed mean the proportion of the total titles that sold greater than 1 unit as that specific format. So, for albums, about 71% of the 6,349 albums sold greater than one CDs. About 38% of all albums sold bundled and unbundled digital units. But for singles, there were more unbundled tracks sold than entire bundled single units, i.e. 24% and 15% respectively. LPs, DVDs, Maxi CDs, 12 inch singles, video tapes, and UMDs are relatively insignificant. However, 34% of the singles sold at least one 12 inch single and 30% of singles sold at least one Maxi CD – greater than both CDs and audio cassettes. For videos, DVDs are more popular than video tapes. However, in this five year period from 2004 to 2008, over 25% of the videos did sell at least one video tape.

Figure 9 compares unit sales for the more popular formats: digital, CDs, and DVDs from 2005 to 2008. Figure 10 compares unit sales for the less popular formats: LPs, cassettes, Maxi CDs, 12 inch singles, video tapes, and UMDs. Although bundled digital unit sales numbers are available in 2004, the remainder of the formats are only available starting in 2005. For these formats, in order to find year-to-date (YTD) sales, release-to-date (RTD) sales for year t are subtracted from year t+1 RTD sales. Because the first year available is 2004, 2003 RTD figures are not available, and 2004 YTD numbers cannot be computed for the titles released before 2003. This is why Figures 9 and 10 only show a four year period.

Table 3: Descriptive Statistics for Albums, Singles, and Videos: Sales Shares of Specific Formats.

All titles are either albums, singles, or videos. Albums can be sold as LPs, cassettes, CDs, DVDs, and bundled digital form. Singles can be sold as cassettes, CDs, Maxi CDs, and 12 Inch singles. Videos can be sold as DVDs, video tapes, and UMDs.

Descriptive Statistics								
	All	Formats:	Albums		Singles		Videos	
	N	%	N	%	N	%	N	%
	7010	100.00%	6349	90.57%	310	4.42%	351	5.01%
Have sold >0 units as:								
LPs	211	3.01%	211	3.32%	N/A	N/A	N/A	N/A
Cassettes	778	11.10%	696	10.96%	82	26.45%	N/A	N/A
CDs	4594	65.53%	4507	70.99%	87	28.06%	N/A	N/A
DVDs	140	2.00%	7	0.11%	N/A	N/A	133	37.89%
Maxi CDs	92	1.31%	N/A	N/A	92	29.68%	N/A	N/A
12 Inches	106	1.51%	N/A	N/A	106	34.19%	N/A	N/A
Video Tapes	90	1.28%	N/A	N/A	N/A	N/A	90	25.64%
UMDs	1	0.01%	N/A	N/A	N/A	N/A	1	0.28%
Digital	2436	34.75%	2389	37.63%	47	15.16%	N/A	N/A
Tracks (Unbundled)	2457	35.05%	2383	37.53%	73	23.55%	N/A	N/A

Looking at Figure 9, DVD sales peak in 2006 but then bottom out by 2008. This follows the pattern of overall unit sales in Figure 1. CD sales do the same, but the yearly changes are much more pronounced. Starting from 5.4 million in 2005 and peaking at nearly 7 million in 2006, CD sales settle to a new low by 2008 at 2.6 million. Bundled digital album sales rise consistently over the years, but are relatively small in volume. Figure 10 tracks the less popular formats over time. Perhaps not surprisingly given consumer technology trends, audio cassette sales plummet from 2005 to 2008. Another implication from Figure 10 is that LP sales are a niche market on the rise: increasing 107% from 2005 to 2008

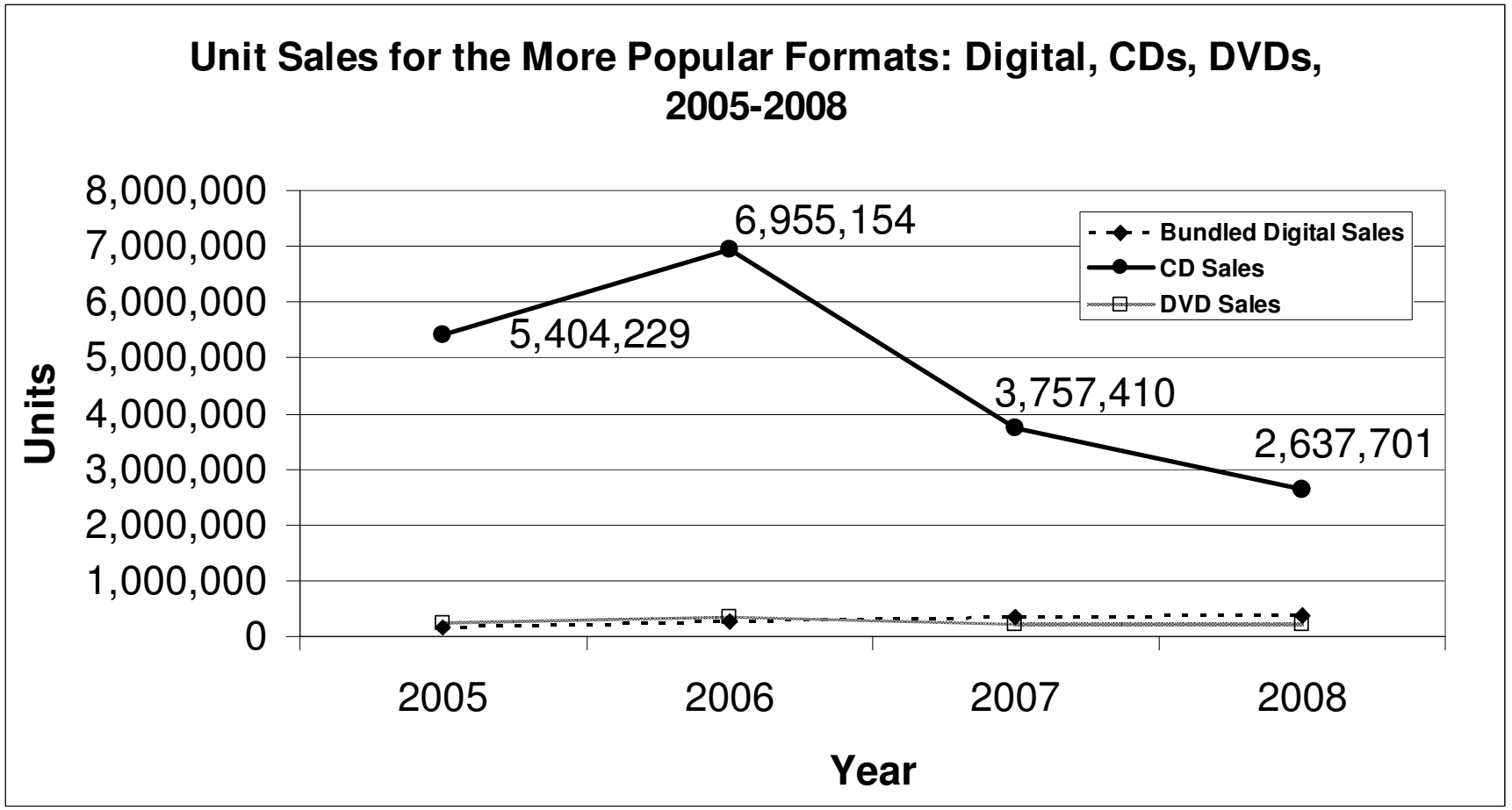

Figure 9: Unit Sales for the More Popular Formats: Digital, CDs, DVDs, 2005-2008.
Bundled digital albums and singles, CDs, and DVDs from 2005 through 2008.

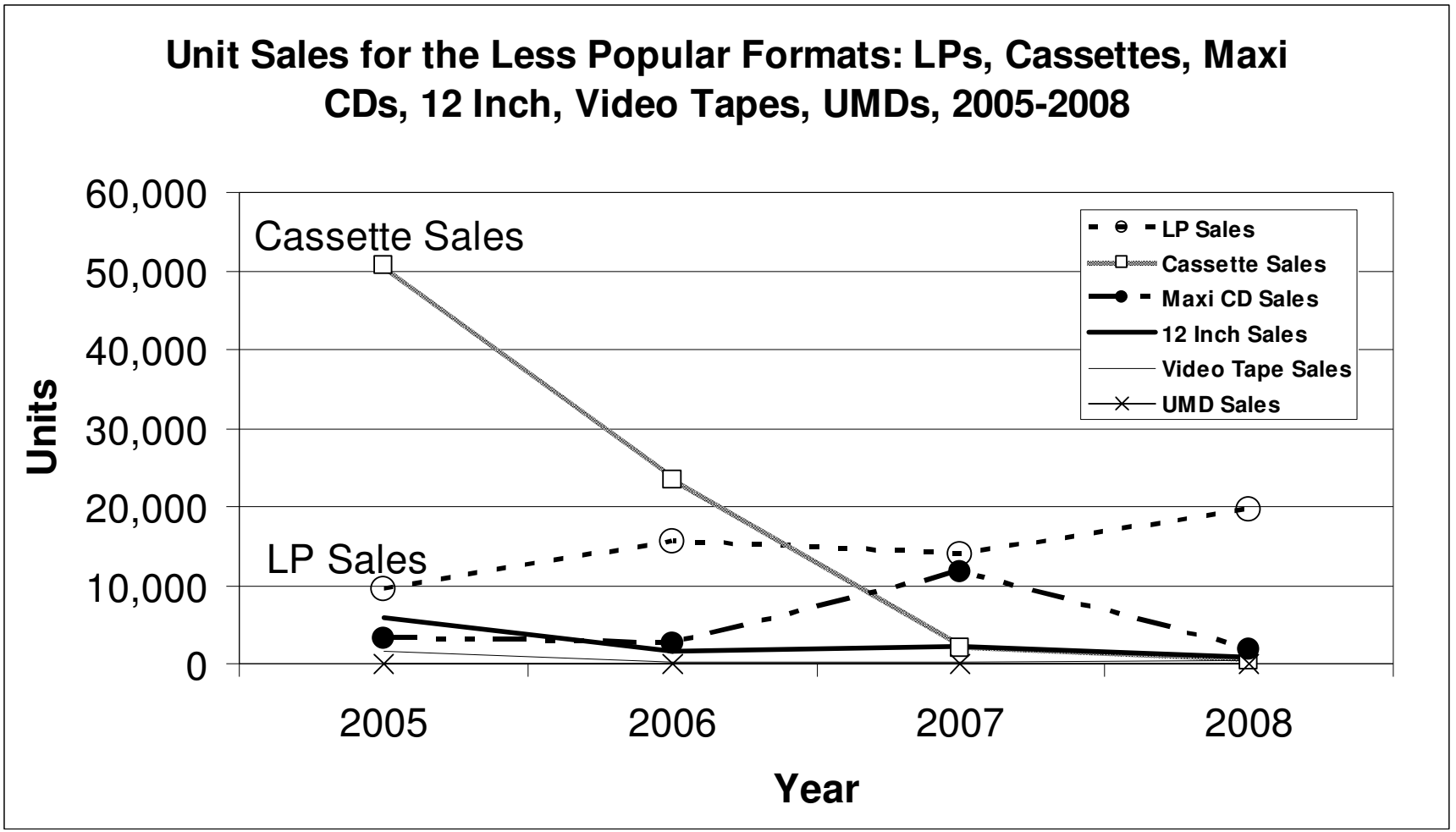

Figure 10: Unit Sales for the Less Popular Formats: LPs, Cassettes, Maxi CDs, 12 Inch, Video Tapes, UMDs, 2005-2008.

LPs, cassettes, Maxi CDs, 12 Inch singles, video tapes, and UMDs.

IV. Empirical Results

This chapter examines the data in three steps. First, relative shifts in the distribution are evaluated. Relative shifts are expressed in terms of percentages, fractions, or ratios without units. This is done by comparing the relative sales shares for different quantiles across time and by comparing Gini coefficients across time. Second, absolute shifts in the sales distribution are examined. Absolute shifts are expressed in terms of sales levels with units. This is done by comparing changes in the level of sales for several summary statistics across time. Third, simple regressions are performed which determine if there are differences in the slope of the sales distribution between years and if these differences are statistically significant.

A. Relative Sales Changes: Is the Pie Sliced Differently?

In this section, the distribution of sales is investigated for artists, titles, and labels for each year of the sample by comparing changes in the proportion of the total sales captured by different percentiles of artists, titles, and labels respectively. To start, relative changes in the distribution of sales are analyzed by quantile. Subsequently, Gini coefficients are compared in the same manner.

Following the analysis of Elberse and Oberholzer-Gee (2008), Table 4 shows sales distribution percentages corresponding to multiple quantiles of artists for each year in the

sample. This shows, for all formats taken together, the percentage of total sales received by each decile of artists, as well as by each percentile above the 90th. These numbers imply that the market becomes more concentrated from 2004 to 2006. This is because each quantile of artists shown from the 60th on up to the 99th account for a smaller percentage of total sales from 2004 to 2006. The top 1% of artists increases until 2006 and is the driver this change. More concentration, in a relative sense, implies the tail of the distribution gets less. This suggests a superstar effect. However, later on the market becomes less concentrated. From 2006 to 2008 for artists there is less concentration because these same quantiles account for an increasing percentage of total sales. Less concentration implies the tail of the distribution gets more. This suggests a long tail effect. Specifically, in 2004, 95% of artists account for 2.65% of total sales. This drops to 1.08% by 2006, but rises again to nearly 2% in 2008. In 2004, 99% of artists account for 17.65% of total sales. This dips to a low of 7.35% in 2006, but rises again to 12.62% by 2008. Figures 11, 12, and 13 illustrate this distributional pattern for artists by graphing the percentage of total sales captured by each quantile listed for each year. This depicts the increasing and then decreasing trend in relative concentration.

Table 5 also shows the percentage of total sales received by each quantile of artist – which is the percent cumulative frequency of unit sales by artist – juxtaposed with the cumulative frequency of unit sales by artist. This is done for ease of comparison. The cumulative frequency is calculated by summing unit sales 'so far' up to a given percentile of artists. This can be considered the cumulative distribution function (CDF). The percent cumulative frequency is calculated by dividing the cumulative frequency at the given percentile by total cumulative unit sales.

Table 4: The Percent of Total Sales Accruing to Different Quantiles of Artists, 2004-2008.

Specifically, this shows the 1st through 9th decile, and then 91st through 99th percentile of artists. For example, the bottom 95% of artists account for 2.64% of sales in 2004, but the bottom 95% of artists account for 1.40% of sales in 2007, and so on.

	% sales	All Formats			
	N=1172	N=1336	N=1575	N=1795	N=2051
% artists	2004	2005	2006	2007	2008
10	0.0000%	0.0000%	0.0000%	0.0000%	0.0000%
20	0.0000%	0.0000%	0.0000%	0.0000%	0.0000%
30	0.0000%	0.0000%	0.0000%	0.0000%	0.0000%
40	0.0000%	0.0000%	0.0003%	0.0003%	0.0004%
50	0.0028%	0.0025%	0.0032%	0.0050%	0.0005%
60	0.0119%	0.0102%	0.0105%	0.0165%	0.0257%
70	0.0404%	0.0331%	0.0300%	0.0459%	0.0693%
80	0.1620%	0.1119%	0.0879%	0.1283%	0.1807%
90	0.8631%	0.5586%	0.3778%	0.4958%	0.6919%
91	1.0540%	0.6916%	0.4509%	0.5913%	0.8237%
92	1.2678%	0.8498%	0.5385%	0.7192%	1.0033%
93	1.5954%	1.0709%	0.6650%	0.8793%	1.2239%
94	2.0274%	1.3823%	0.8280%	1.1018%	1.5393%
95	2.6468%	1.8380%	1.0749%	1.4057%	1.9449%
96	3.7797%	2.6744%	1.4625%	1.8441%	2.6160%
97	5.9164%	3.9328%	2.1797%	2.7226%	3.8100%
98	10.7007%	6.2677%	3.5702%	5.0546%	6.2428%
99	17.6450%	11.5334%	7.3497%	10.9743%	12.6179%
100	100.0000%	100.0000%	100.0000%	100.0000%	100.0000%

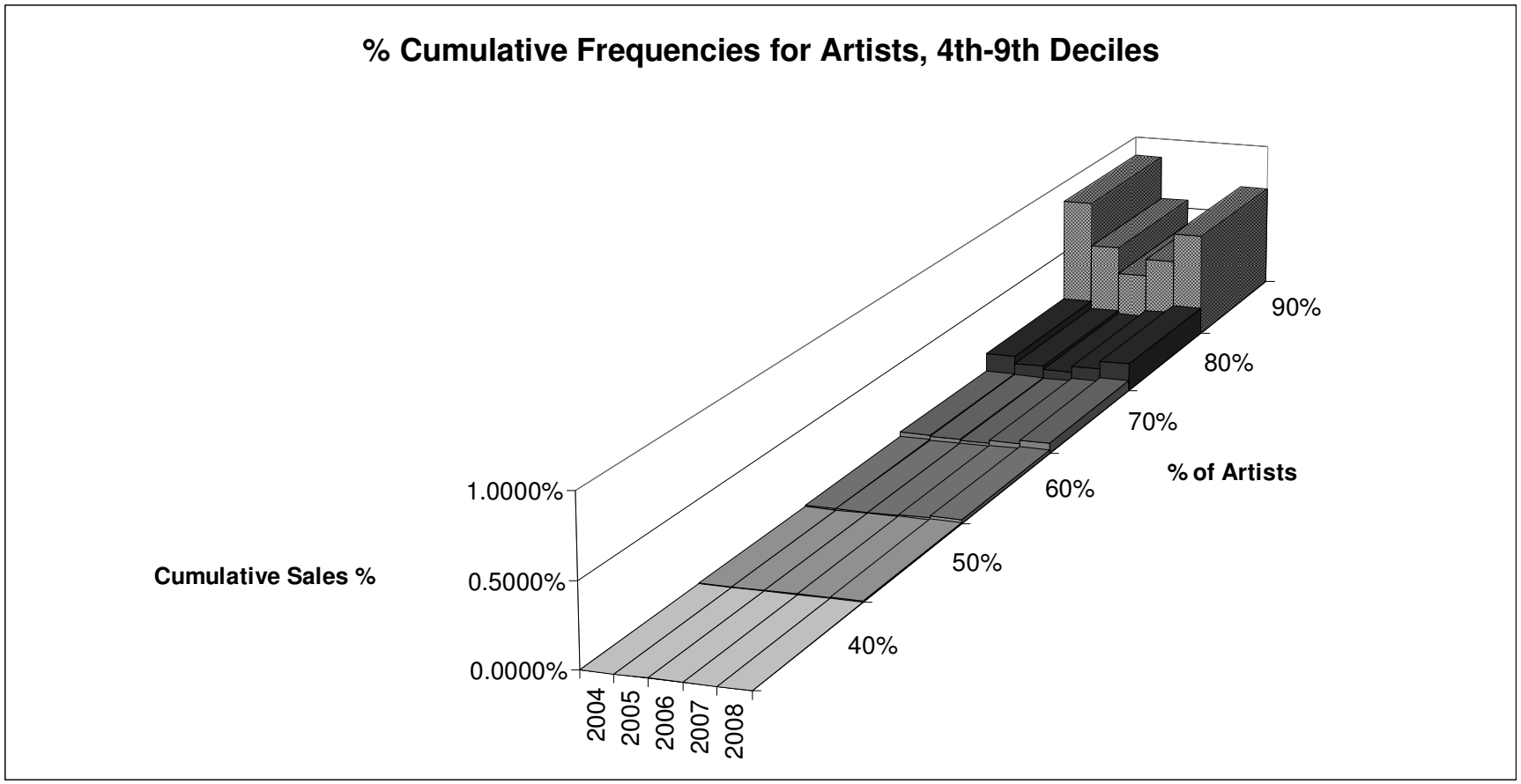

Figure 11: Percent Cumulative Frequencies – Percent of Total Sales Accruing to 4th through 9th Deciles of Artist Unit Sales, 2004-2008.

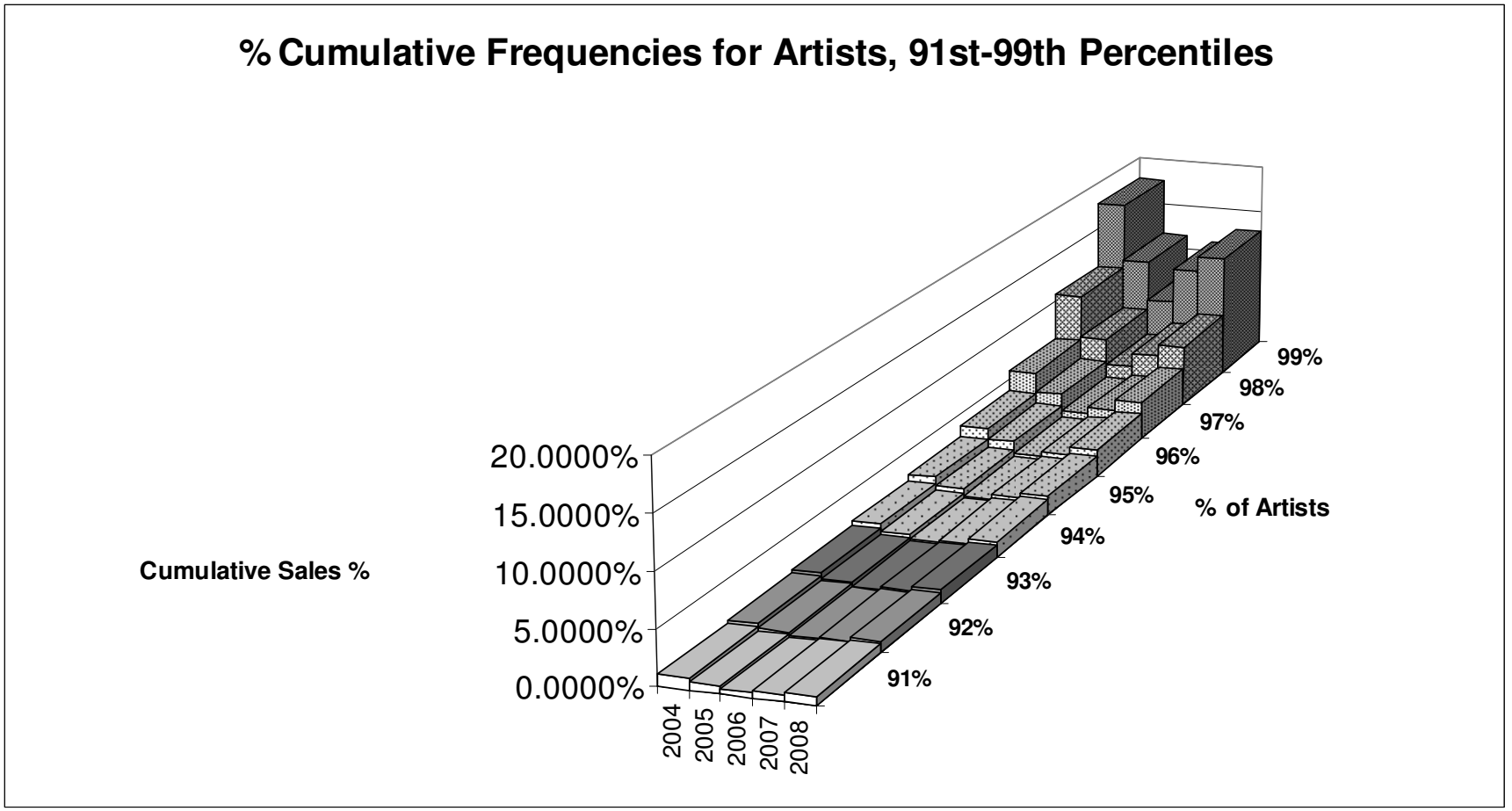

Figure 12: Percent Cumulative Frequencies – Percent of Total Sales Accruing to 91st through 99th Percentiles of Artist Unit Sales, 2004-2008.

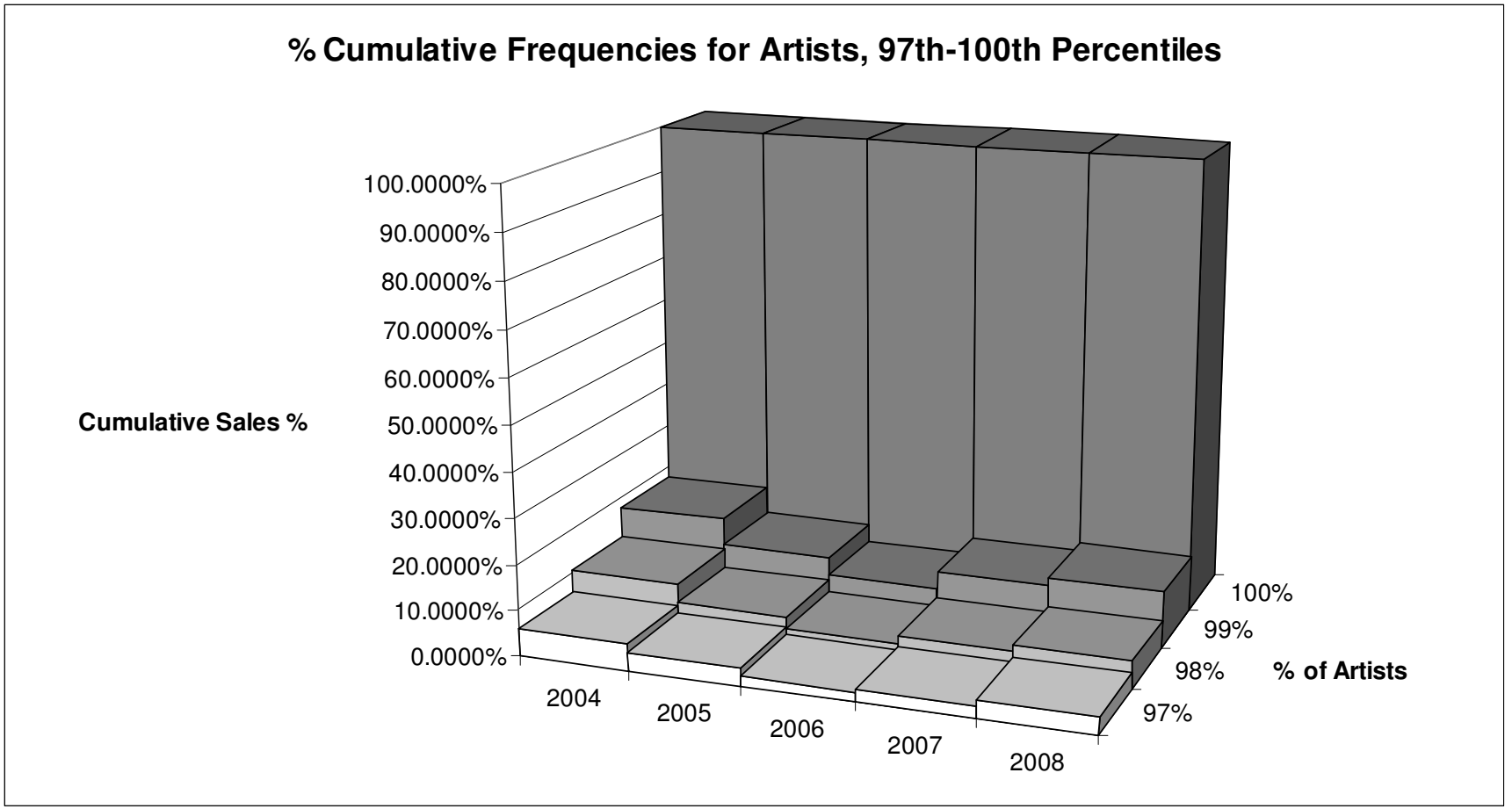

Figure 13: Percent Cumulative Frequencies – Percent of Total Sales Accruing to 97th through 100th Percentiles of Artist Unit Sales, 2004-2008.

Table 5: Cumulative Frequencies and % Cumulative Frequencies for Artists, 2004-2008.

As with Table 4, this table also shows the percent of total sales accruing to different quantiles of artists, which is the % cumulative frequency of unit sales. For ease of comparison, this is shown here alongside the cumulative frequency of unit sales for the quantile of artists shown in the first column.

Unit Sales - Cumulative Frequencies and % Cumulative Frequencies for Artists										
	N=1172		N=1336		N=1575		N=1795		N=2051	
	2004		2005		2006		2007		2008	
% Artists	Cum. Freq.	% Cum. Freq.	Cum. Freq.	% Cum. Freq.	Cum. Freq.	% Cum. Freq.	Cum. Freq.	% Cum. Freq.	Cum. Freq.	% Cum. Freq.
10%	0	0.0000%	0	0.0000%	0	0.0000%	0	0.0000%	0	0.0000%
20%	0	0.0000%	0	0.0000%	0	0.0000%	0	0.0000%	0	0.0000%
30%	0	0.0000%	0	0.0000%	0	0.0000%	0	0.0000%	0	0.0000%
40%	0	0.0000%	0	0.0000%	25	0.0003%	11	0.0003%	14	0.0004%
50%	140	0.0028%	140	0.0025%	239	0.0032%	219	0.0050%	253	0.0077%
60%	600	0.0119%	581	0.0102%	775	0.0105%	723	0.0165%	840	0.0257%
70%	2,033	0.0404%	1,889	0.0331%	2,210	0.0300%	2,007	0.0459%	2,267	0.0693%
80%	8,160	0.1620%	6,388	0.1119%	6,489	0.0879%	5,612	0.1283%	5,909	0.1807%
90%	43,468	0.8631%	31,900	0.5586%	27,877	0.3778%	21,688	0.4958%	22,625	0.6919%
91%	53,084	1.0540%	39,491	0.6916%	33,272	0.4509%	25,867	0.5913%	26,934	0.8237%
92%	63,853	1.2678%	48,526	0.8498%	39,738	0.5385%	31,462	0.7192%	32,807	1.0033%
93%	80,350	1.5954%	61,148	1.0709%	49,067	0.6650%	38,464	0.8793%	40,020	1.2239%
94%	102,108	2.0274%	78,932	1.3823%	61,099	0.8280%	48,198	1.1018%	50,333	1.5393%
95%	133,302	2.6468%	104,952	1.8380%	79,312	1.0749%	61,494	1.4057%	63,596	1.9449%
96%	190,357	3.7797%	152,715	2.6744%	107,914	1.4625%	80,672	1.8441%	85,539	2.6160%
97%	297,972	5.9164%	224,569	3.9328%	160,838	2.1797%	119,101	2.7226%	124,580	3.8100%
98%	538,924	10.7007%	357,902	6.2677%	263,436	3.5702%	221,117	5.0546%	204,127	6.2428%
99%	888,668	17.6450%	658,580	11.5334%	542,320	7.3497%	480,077	10.9743%	412,583	12.6179%
100%	5,036,364	100.0000%	5,710,220	100.0000%	7,378,794	100.0000%	4,374,567	100.0000%	3,269,817	100.0000%

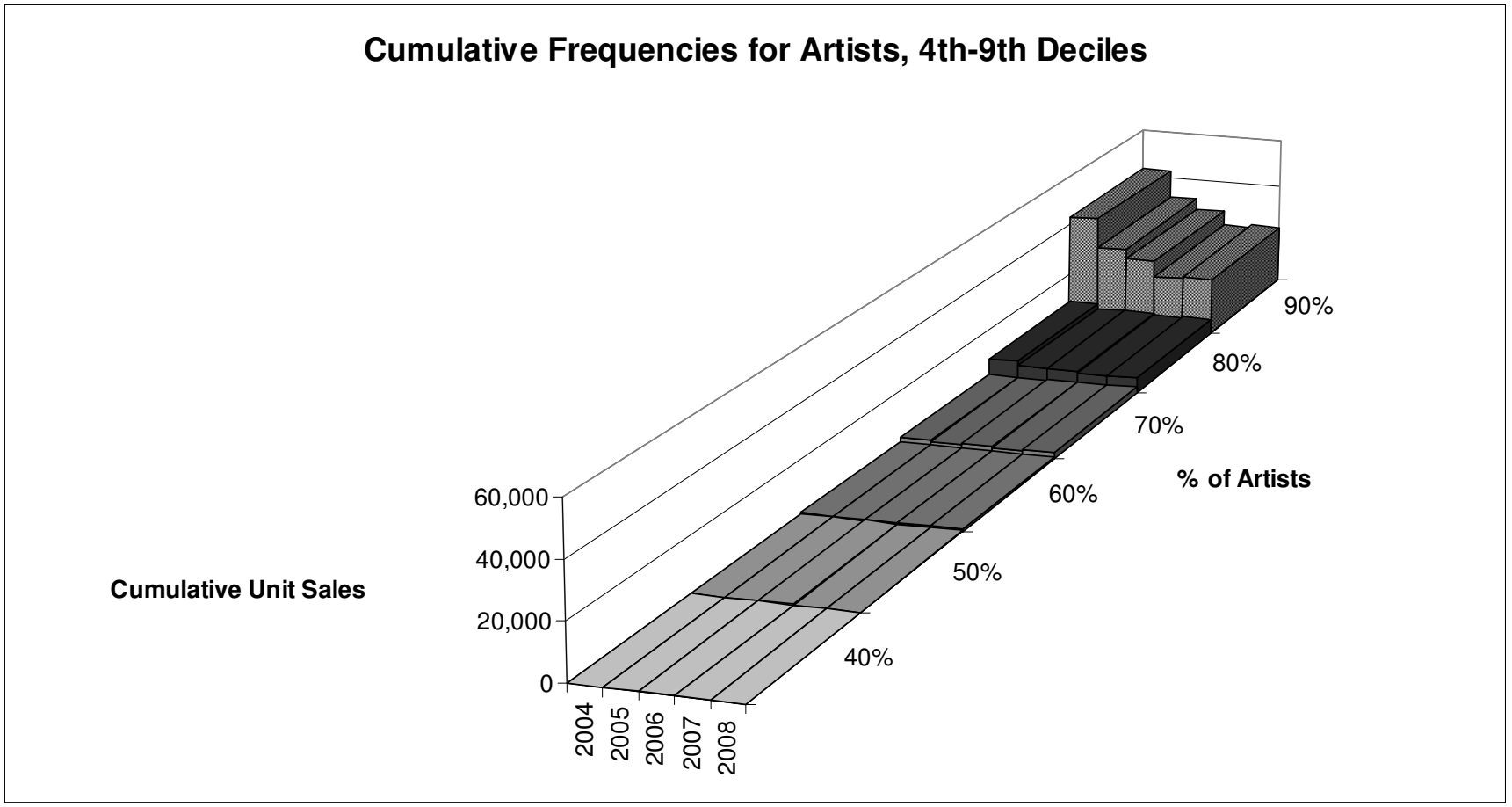

Figure 14: Cumulative Frequencies for 4th through 9th Deciles of Artist Sales, 2004-2008.

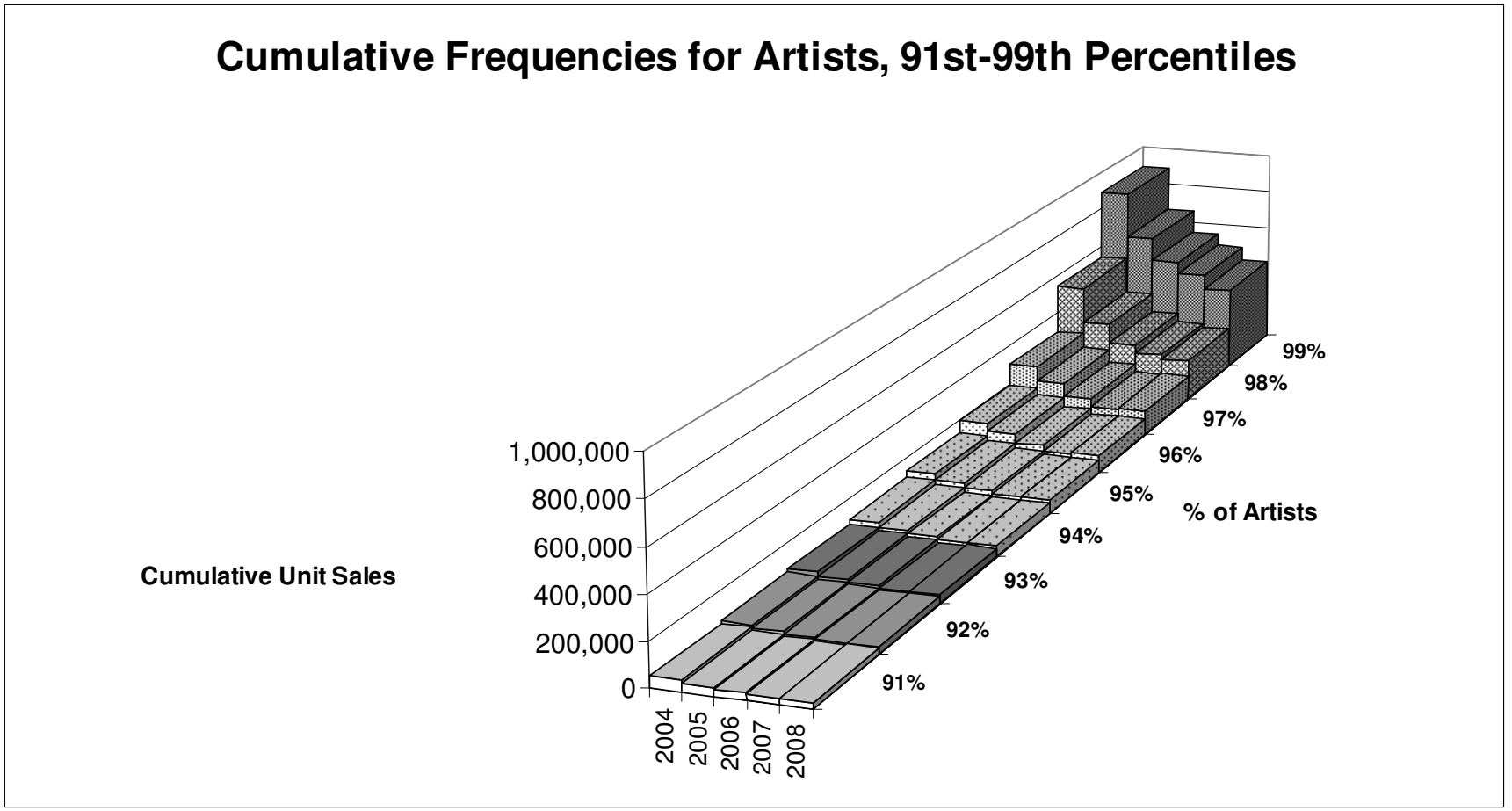

Figure 15: Cumulative Frequencies for 91st through 99th Percentiles of Artist Unit Sales, 2004-2008.

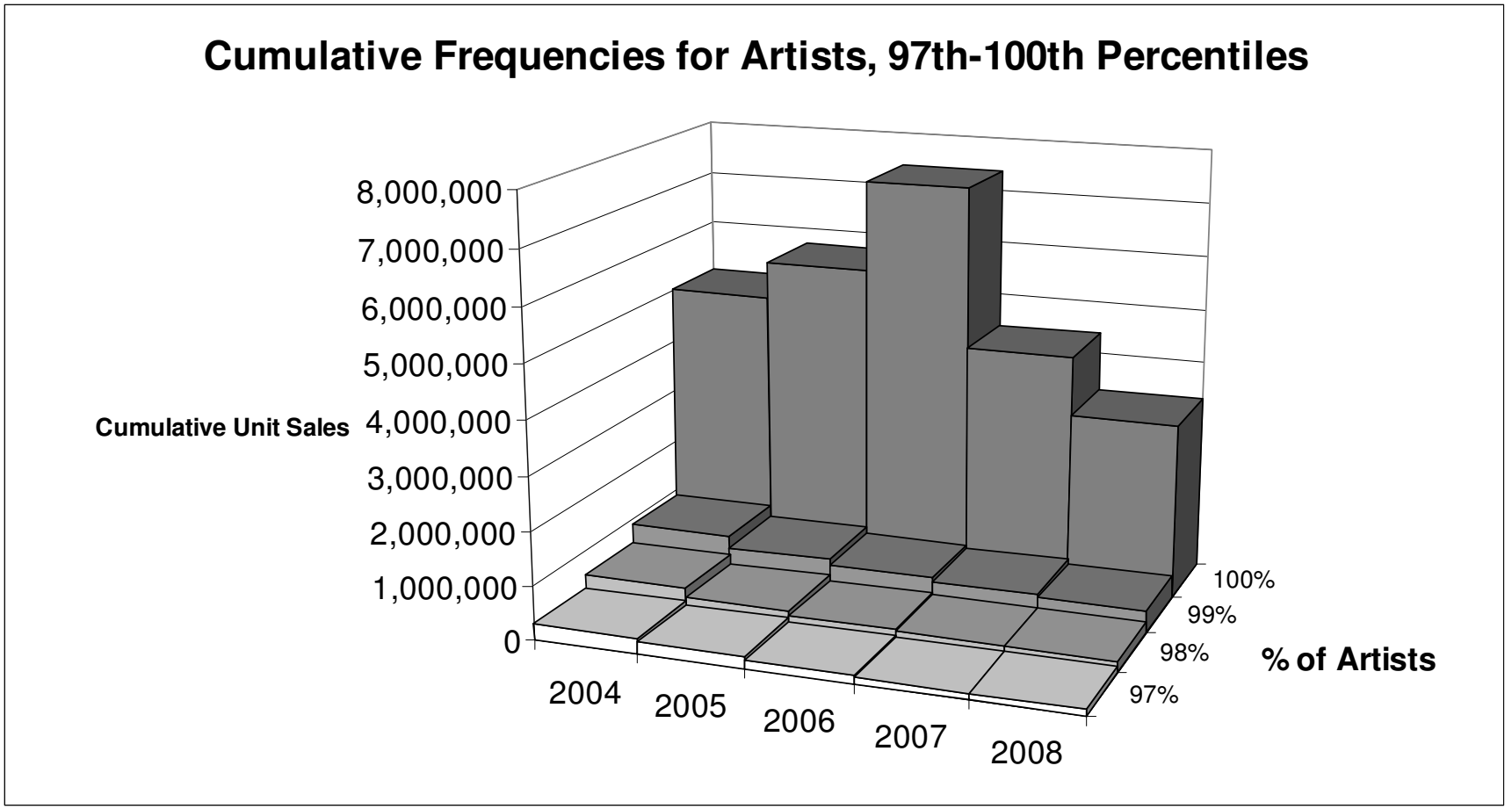

Figure 16: Cumulative Frequencies for 97th through 100th Percentiles of Artist Unit Sales, 2004-2008.

Table 6 also displays sales distributions by quantile, but for individual titles. Table 6 lists the percentage of total sales received by each decile of titles, as well as each percentile above the 90th, again, for all formats. Titles exhibit a similar pattern as albums with increasing concentration from 2004 through 2006 and decreasing concentration from 2006 through 2008. Specifically, in 2004, 95% of titles account for 11.255% of total sales. This drops to 5.51% by 2006, but rises again to nearly 7.2% in 2008. In 2004, 99% of artists account for 41% of total sales. This dips to a low of 25% in 2006, but rises to 30% by 2008. The top 1% of titles is the driver of these changes in relative shares. Figures 17, 18, and 19 illustrate this distributional pattern for titles. Again, there seems to be a superstar effect from 2004 through 2006 and a long tail effect from 2006 through 2008.

Table 7 also shows the percentage of total sales received by each quantile of title – which is the percent cumulative frequency of unit sales by title – along with the cumulative frequency of unit sales by title. This is done for ease of comparison of the CDF with the percent cumulative frequency. These figures are arrived at analogous to Table 5.

Table 6: The Percent of Total Sales Accruing to Different Quantiles of Titles, 2004-2008.

Specifically, this shows the 1st through 9th decile, and then 91st through 99th percentile of titles.

	% sales	All Formats			
	N=3882	N=4468	N=5201	N=6088	N=7010
% titles	2004	2005	2006	2007	2008
10	0.0000%	0.0000%	0.0000%	0.0000%	0.0000%
20	0.0000%	0.0000%	0.0000%	0.0000%	0.0000%
30	0.0000%	0.0000%	0.0000%	0.0000%	0.0000%
40	0.0025%	0.0005%	0.0004%	0.0000%	0.0000%
50	0.0201%	0.0121%	0.0100%	0.0090%	0.0144%
60	0.0791%	0.0521%	0.0417%	0.0441%	0.0648%
70	0.2576%	0.1760%	0.1357%	0.1534%	0.2158%
80	0.8794%	0.6150%	0.4538%	0.5074%	0.6872%
90	4.1325%	2.8391%	1.9366%	2.2060%	2.7330%
91	4.9975%	3.4378%	2.3379%	2.6875%	3.2231%
92	6.0387%	4.2001%	2.8611%	3.2963%	3.8625%
93	7.3461%	5.1506%	3.5185%	4.0931%	4.6680%
94	9.0086%	6.4371%	4.3644%	5.1777%	5.7135%
95	11.2480%	8.1419%	5.5126%	6.6698%	7.1970%
96	14.3714%	10.4874%	7.2230%	8.7483%	9.3140%
97	18.8880%	13.8398%	10.0046%	11.8003%	12.4021%
98	26.1212%	19.0523%	14.6091%	17.3723%	18.0411%
99	40.9483%	29.4064%	24.9872%	28.8969%	29.8275%
100	100.0000%	100.0000%	100.0000%	100.0000%	100.0000%

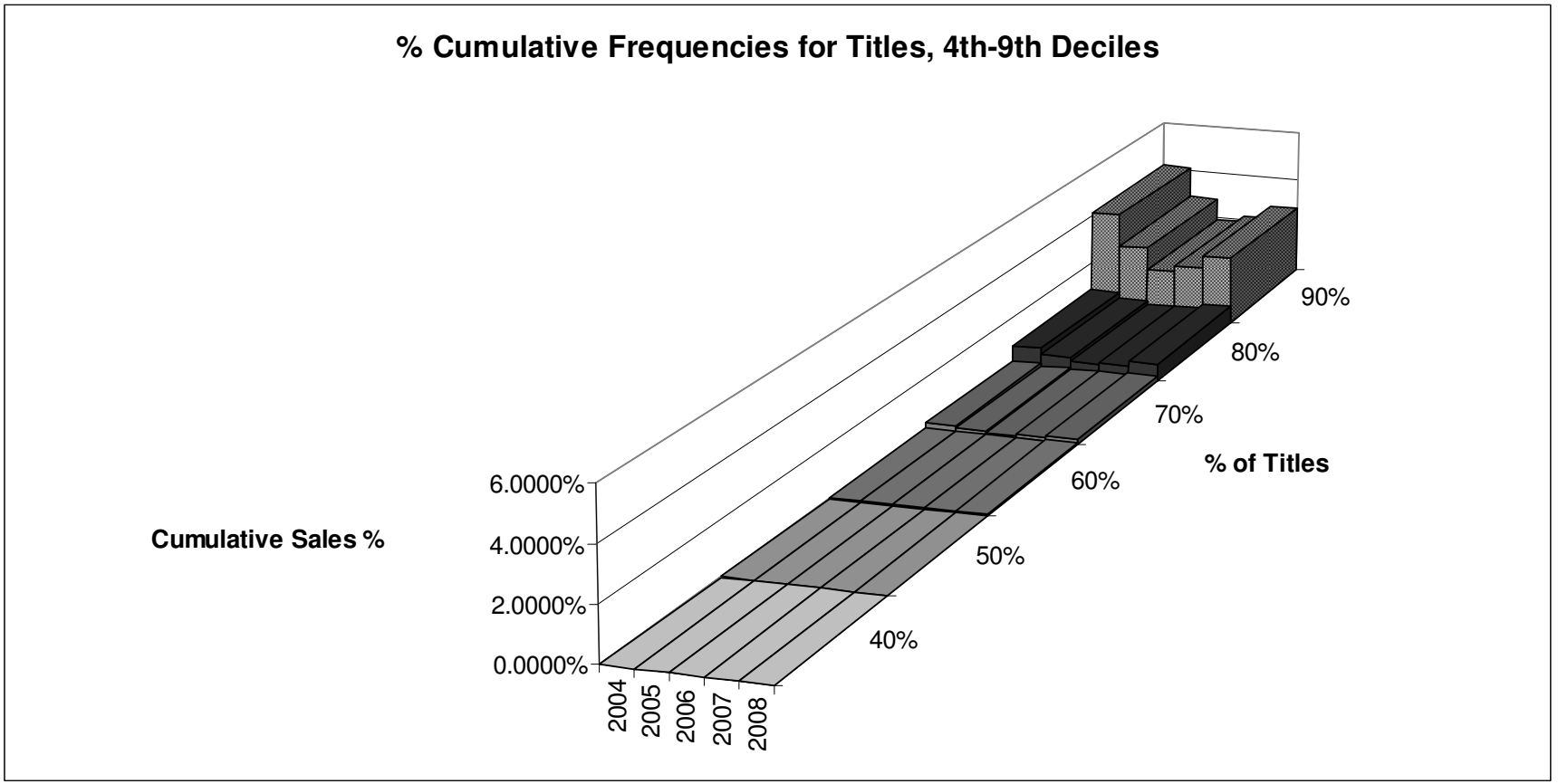

Figure 17: Percent Cumulative Frequencies – Percent of Total Sales Accruing to 4th through 9th Deciles of Title Unit Sales, 2004-2008.

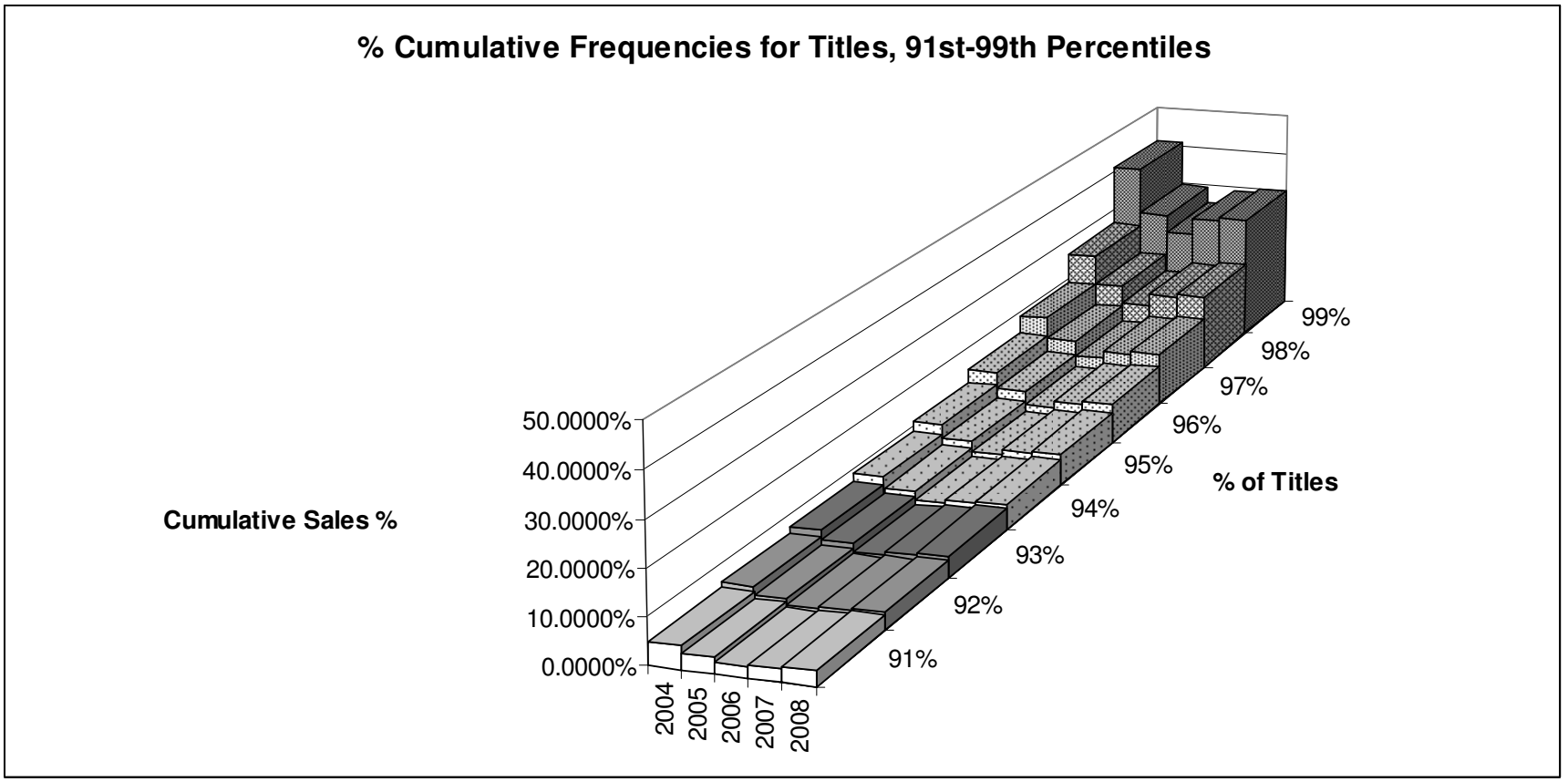

Figure 18: Percent Cumulative Frequencies – Percent of Total Sales Accruing to 91st through 99th Percentiles of Title Unit Sales, 2004-2008.

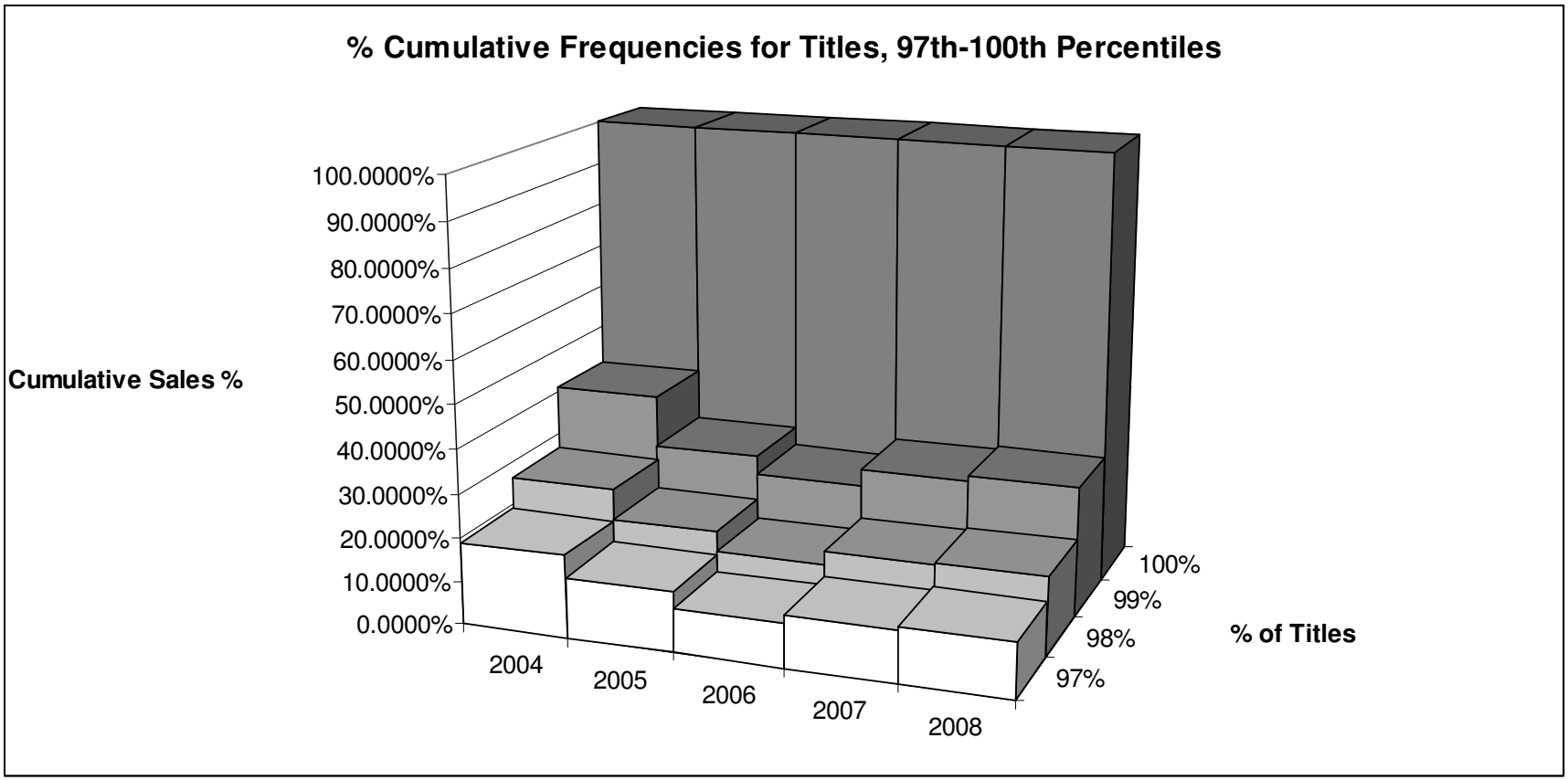

Figure 19: Percent Cumulative Frequencies – Percent of Total Sales Accruing to 97th through 100th Percentiles of Title Unit Sales, 2004-2008.

Table 7: Cumulative Frequencies and % Cumulative Frequencies for Titles, 2004-2008.

As with Table 6, this table also shows the percent of total sales accruing to different quantiles of titles, which is the % cumulative frequency of unit sales. For ease of comparison, this is shown here alongside the cumulative frequency of unit sales for the quantiles of titles shown in the first column.

Unit Sales - Cumulative Frequencies and % Cumulative Frequencies for Titles										
	N=3882		N=4468		N=5201		N=6088		N=7010	
	2004		2005		2006		2007		2008	
% Titles	Cum. Freq.	% Cum. Freq.	Cum. Freq.	% Cum. Freq.	Cum. Freq.	% Cum. Freq.	Cum. Freq.	% Cum. Freq.	Cum. Freq.	% Cum. Freq.
10%	0	0.0000%	0	0.0000%	0	0.0000%	0	0.0000%	0	0.0000%
20%	0	0.0000%	0	0.0000%	0	0.0000%	0	0.0000%	0	0.0000%
30%	0	0.0000%	0	0.0000%	0	0.0000%	0	0.0000%	0	0.0000%
40%	125	0.0025%	27	0.0005%	30	0.0004%	0	0.0000%	0	0.0000%
50%	1,013	0.0201%	691	0.0121%	741	0.0100%	395	0.0090%	471	0.0144%
60%	3,986	0.0791%	2,975	0.0521%	3,074	0.0417%	1,930	0.0441%	2,120	0.0648%
70%	12,973	0.2576%	10,051	0.1760%	10,016	0.1357%	6,712	0.1534%	7,057	0.2158%
80%	44,292	0.8794%	35,115	0.6150%	33,486	0.4538%	22,198	0.5074%	22,470	0.6872%
90%	208,127	4.1325%	162,120	2.8391%	142,896	1.9366%	96,501	2.2060%	89,364	2.7330%
91%	251,691	4.9975%	196,304	3.4378%	172,512	2.3379%	117,568	2.6875%	105,389	3.2231%
92%	304,130	6.0387%	239,833	4.2001%	211,116	2.8611%	144,197	3.2963%	126,298	3.8625%
93%	369,976	7.3461%	294,110	5.1506%	259,622	3.5185%	179,057	4.0931%	152,636	4.6680%
94%	453,708	9.0086%	367,575	6.4371%	322,041	4.3644%	226,500	5.1777%	186,823	5.7135%
95%	566,488	11.2480%	464,920	8.1419%	406,767	5.5126%	291,777	6.6698%	235,329	7.1970%
96%	723,797	14.3714%	598,851	10.4874%	532,971	7.2230%	382,702	8.7483%	304,551	9.3140%
97%	951,268	18.8880%	790,285	13.8398%	738,219	10.0046%	516,213	11.8003%	405,527	12.4021%
98%	1,315,561	26.1212%	1,087,926	19.0523%	1,077,974	14.6091%	759,961	17.3723%	589,911	18.0411%
99%	2,062,305	40.9483%	1,679,168	29.4064%	1,843,751	24.9872%	1,264,114	28.8969%	975,306	29.8275%
100%	5,036,364	100.0000%	5,710,220	100.0000%	7,378,794	100.0000%	4,374,568	100.0000%	3,269,826	100.0000%

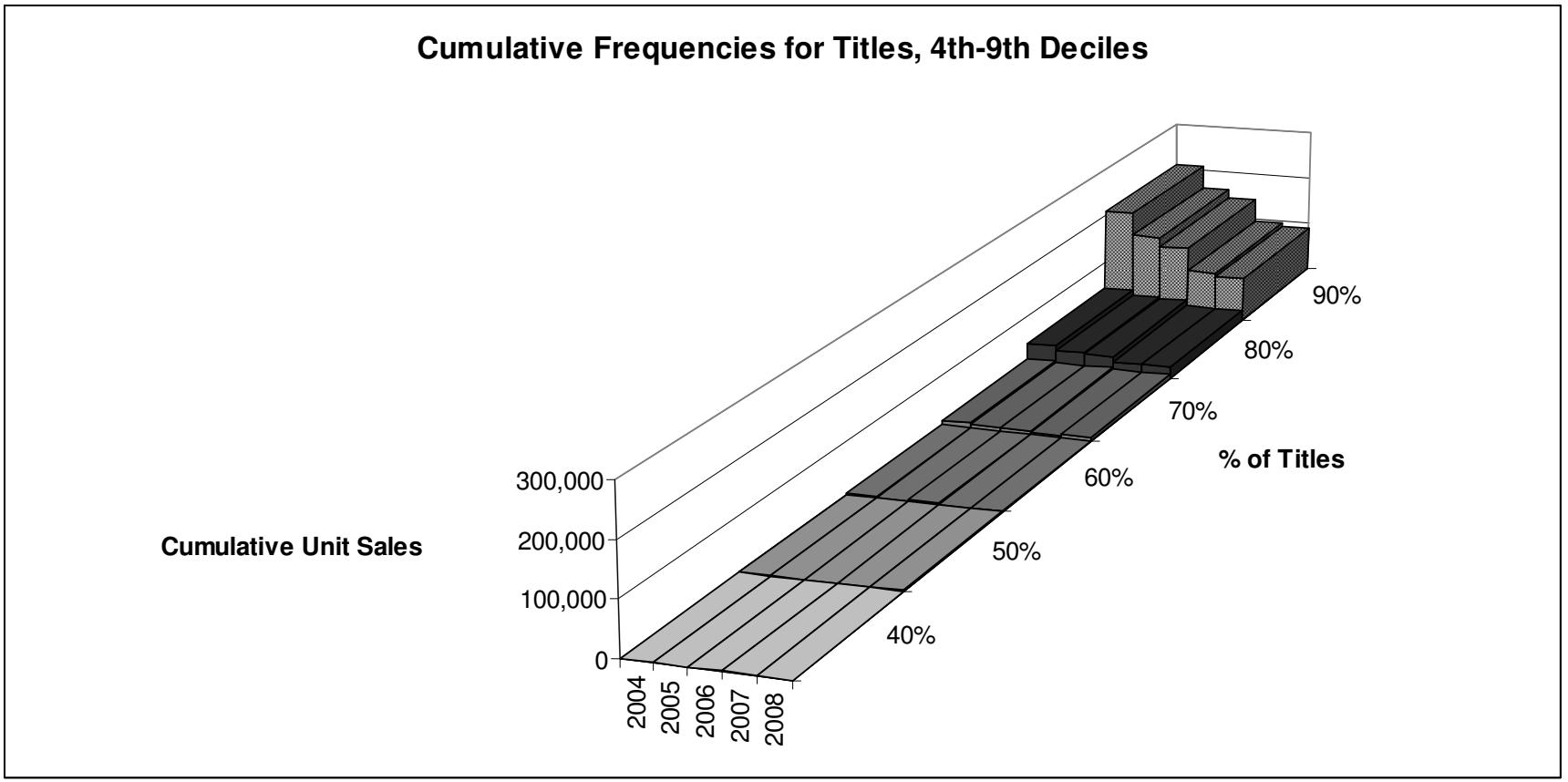

Figure 20: Cumulative Frequencies for 4th through 9th Deciles of Title Sales, 2004-2008.

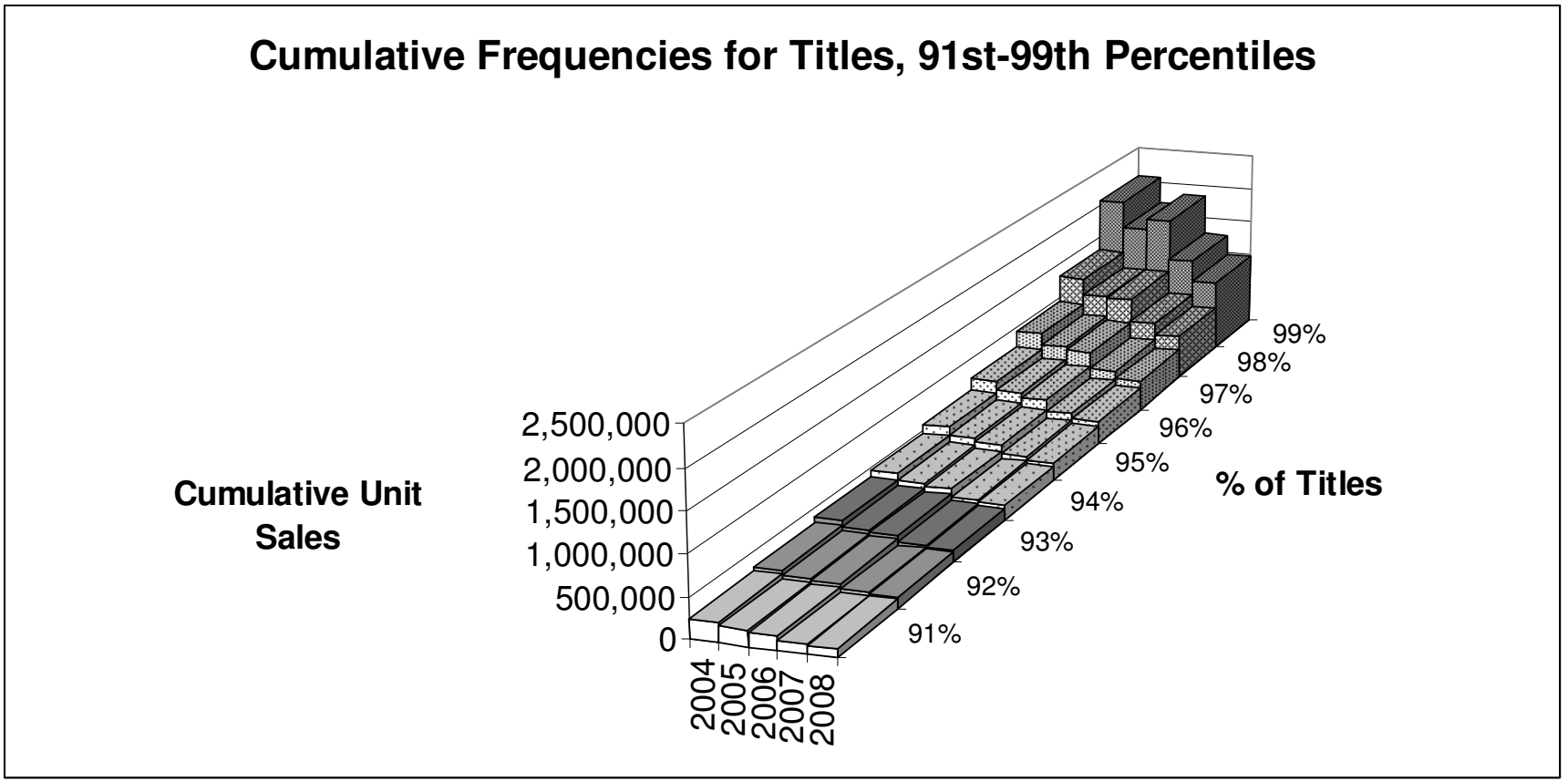

Figure 21: Cumulative Frequencies for 91st through 99th Percentiles of Title Unit Sales, 2004-2008.

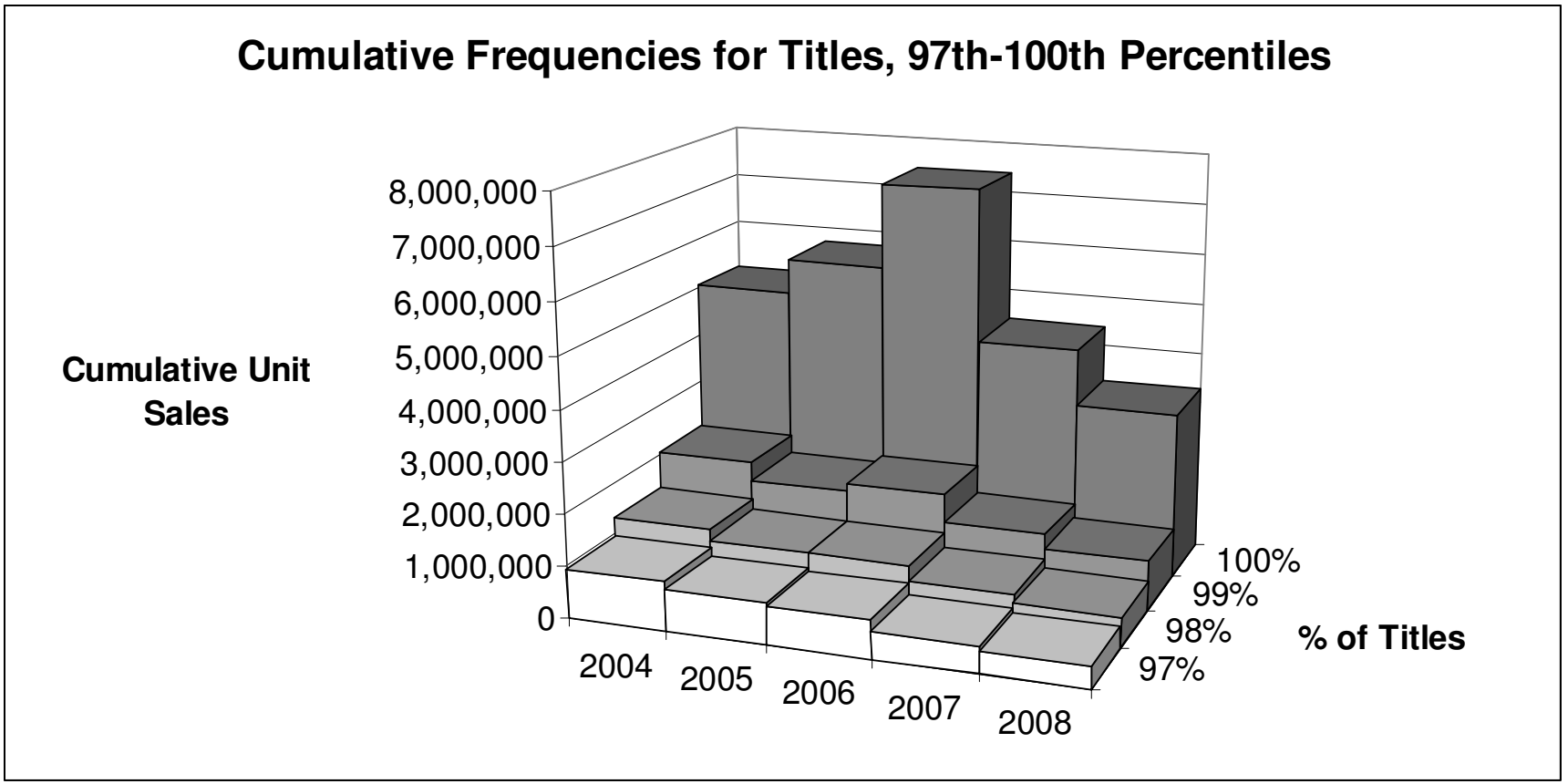

Figure 22: Cumulative Frequencies for 97th through 100th Percentiles of Title Unit Sales, 2004-2008.

Labels show a similar pattern as artists and titles. This can be seen in Table 8 and Figures 23, 24, and 25. The 91st through 99th percentiles, however, do not see the same 'rebound' in sales share in the final two years – 2007 and 2008 – as artists and titles do. The top 1% of labels retain their share of relative sales, so there is a much smaller long tail effect in those last two years judging by the concentration of relative sales by label compared to artists and titles.

Once more, Table 9 also shows the percentage of total sales received by each quantile of label – which is the percent cumulative frequency of unit sales by label – compared with the cumulative frequency of unit sales by label. This is done for ease of comparison of the CDF with the percent cumulative frequency. These figures are arrived at analogous to Table 5.

Table 8: The Percent of Total Sales Accruing to Different Quantiles of Labels, 2004-2008.

Specifically, this shows the 1st through 9th decile, and then 91st through 99th percentile of labels.

	% sales	All Formats			
	N=1104	N=1286	N=1468	N=1643	N=1836
% labels	2004	2005	2006	2007	2008
10	0.0000%	0.0000%	0.0000%	0.0000%	0.0000%
20	0.0000%	0.0000%	0.0000%	0.0000%	0.0000%
30	0.0000%	0.0000%	0.0000%	0.0000%	0.0000%
40	0.0024%	0.0012%	0.0012%	0.0019%	0.0034%
50	0.0111%	0.0081%	0.0075%	0.0091%	0.0155%
60	0.0374%	0.0268%	0.0255%	0.0308%	0.0498%
70	0.1249%	0.0824%	0.0722%	0.0922%	0.1461%
80	0.3895%	0.2575%	0.2022%	0.2780%	0.4311%
90	1.8659%	1.1275%	0.8021%	1.0505%	1.5374%
91	2.2824%	1.4062%	0.9799%	1.2601%	1.8401%
92	2.8616%	1.8358%	1.2251%	1.5751%	2.2301%
93	3.6981%	2.4219%	1.5651%	2.0016%	2.7465%
94	5.0049%	3.3372%	2.1542%	2.6403%	3.4718%
95	7.0933%	4.5987%	3.0464%	3.6507%	4.4454%
96	10.2252%	6.4880%	4.3319%	5.3878%	5.9159%
97	14.3748%	9.1767%	6.8735%	8.3861%	8.3038%
98	19.9591%	14.0426%	11.8891%	13.7415%	13.2370%
99	32.2171%	23.8760%	21.5944%	24.5173%	24.1479%
100	100.0000%	100.0000%	100.0000%	100.0000%	100.0000%

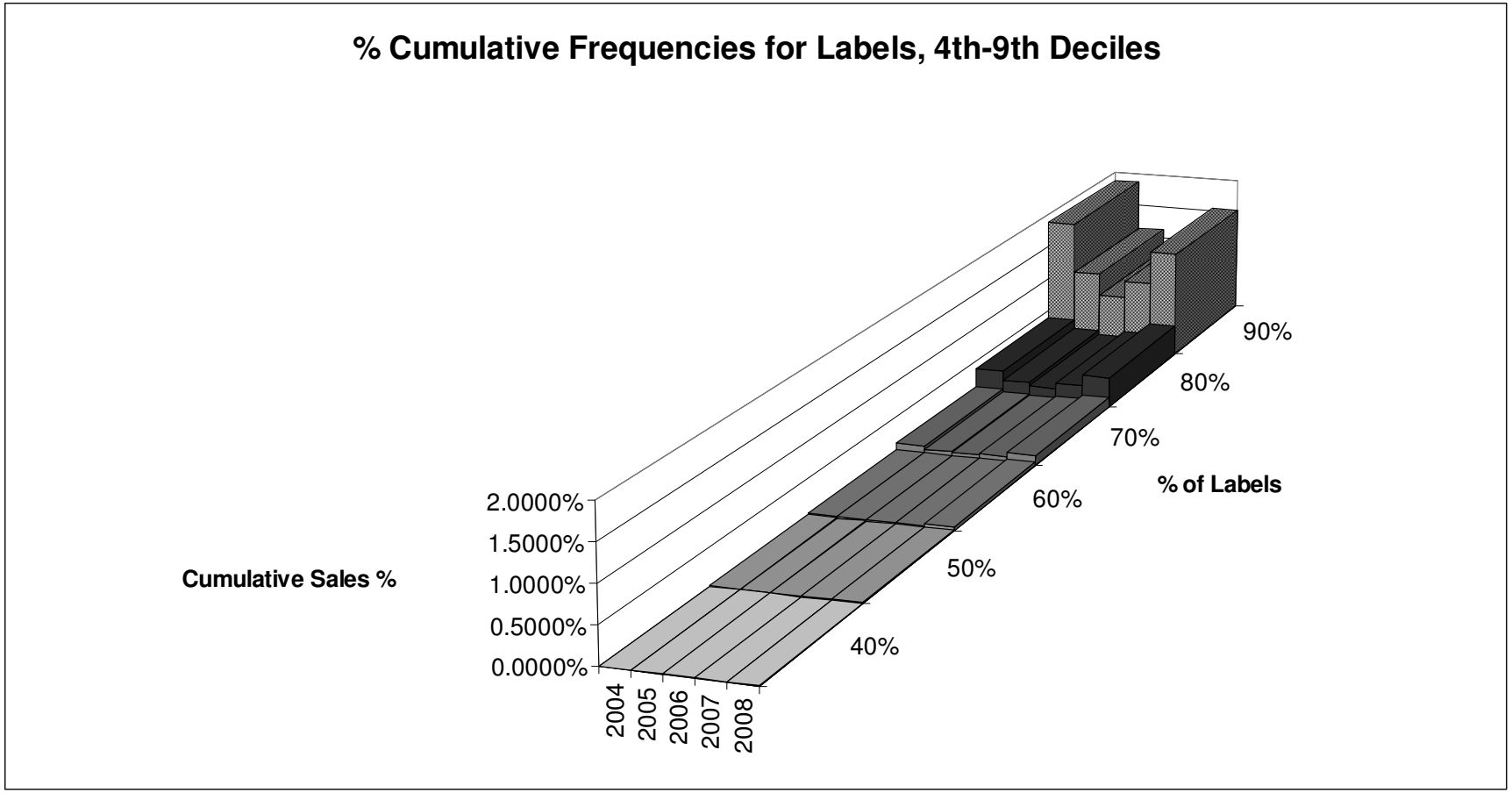

Figure 23: Percent Cumulative Frequencies – Percent of Total Sales Accruing to 4th through 9th Deciles of Label Unit Sales, 2004-2008.

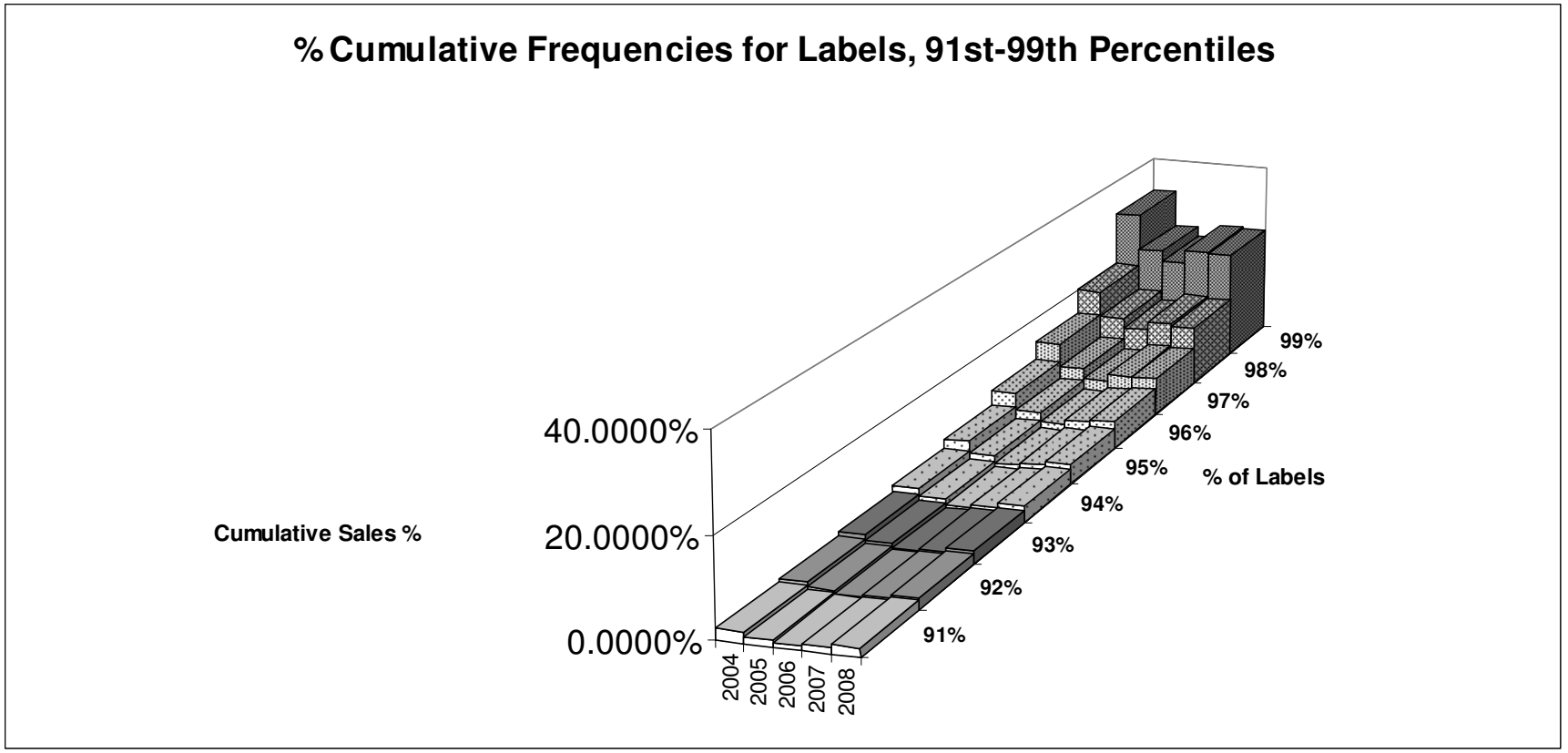

Figure 24: Percent Cumulative Frequencies – Percent of Total Sales Accruing to 91st through 99th Percentiles of Label Unit Sales, 2004-2008.

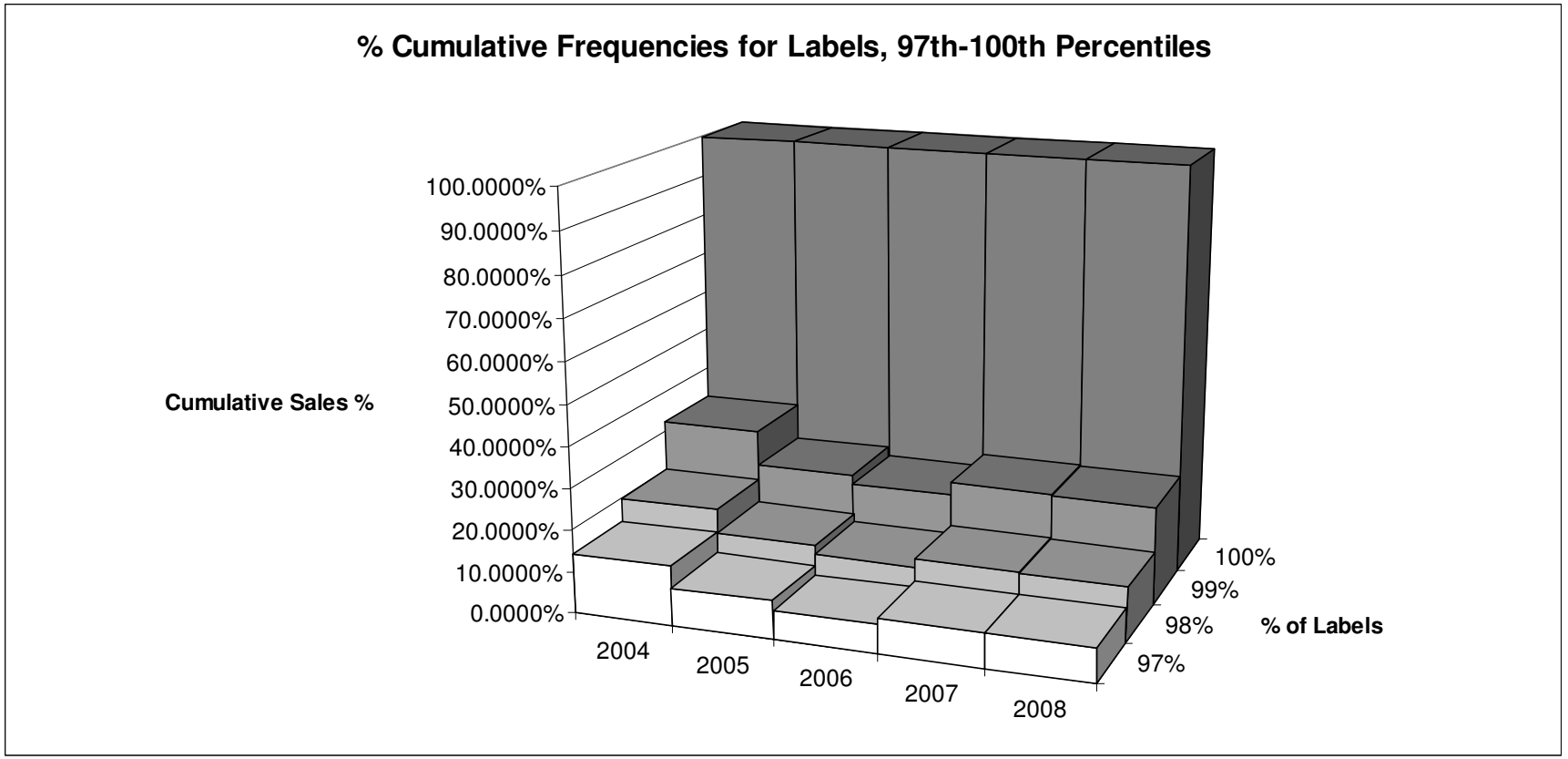

Figure 25: Percent Cumulative Frequencies – Percent of Total Sales Accruing to 97th through 100th Percentiles of Title Unit Sales, 2004-2008.

Table 9: Cumulative Frequencies and % Cumulative Frequencies for Labels, 2004-2008.

As with Table 8, this table also shows the percent of total sales accruing to different quantiles of labels, which is the % cumulative frequency of unit sales. For ease of comparison, this is shown here alongside the cumulative frequency of unit sales for the quantiles of labels shown in the first column.

Unit Sales - Cumulative Frequencies and % Cumulative Frequencies for Labels										
	N=1104		N=1286		N=1468		N=1643		N=1836	
	2004		2005		2006		2007		2008	
% Labels	Cum. Freq.	% Cum. Freq.	Cum. Freq.	% Cum. Freq.	Cum. Freq.	% Cum. Freq.	Cum. Freq.	% Cum. Freq.	Cum. Freq.	% Cum. Freq.
10%	0	0.0000%	0	0.0000%	0	0.0000%	0	0.0000%	0	0.0000%
20%	0	0.0000%	0	0.0000%	0	0.0000%	0	0.0000%	0	0.0000%
30%	0	0.0000%	0	0.0000%	0	0.0000%	0	0.0000%	0	0.0000%
40%	120	0.0024%	67	0.0012%	89	0.0012%	82	0.0019%	112	0.0034%
50%	561	0.0111%	460	0.0081%	553	0.0075%	398	0.0091%	507	0.0155%
60%	1,885	0.0374%	1,531	0.0268%	1,883	0.0255%	1,346	0.0308%	1,628	0.0498%
70%	6,292	0.1249%	4,708	0.0824%	5,328	0.0722%	4,032	0.0922%	4,778	0.1461%
80%	19,619	0.3895%	14,705	0.2575%	14,920	0.2022%	12,163	0.2780%	14,096	0.4311%
90%	93,976	1.8659%	64,384	1.1275%	59,185	0.8021%	45,957	1.0505%	50,270	1.5374%
91%	114,951	2.2824%	80,295	1.4062%	72,303	0.9799%	55,122	1.2601%	60,167	1.8401%
92%	144,122	2.8616%	104,826	1.8358%	90,397	1.2251%	68,904	1.5751%	72,920	2.2301%
93%	186,249	3.6981%	138,293	2.4219%	115,486	1.5651%	87,563	2.0016%	89,807	2.7465%
94%	252,066	5.0049%	190,564	3.3372%	158,957	2.1542%	115,502	2.6403%	113,520	3.4718%
95%	357,244	7.0933%	262,596	4.5987%	224,789	3.0464%	159,701	3.6507%	145,356	4.4454%
96%	514,977	10.2252%	370,481	6.4880%	319,642	4.3319%	235,693	5.3878%	193,438	5.9159%
97%	723,965	14.3748%	524,011	9.1767%	507,184	6.8735%	366,857	8.3861%	271,519	8.3038%
98%	1,005,212	19.9591%	801,865	14.0426%	877,271	11.8891%	601,133	13.7415%	432,825	13.2370%
99%	1,622,571	32.2171%	1,363,369	23.8760%	1,593,405	21.5944%	1,072,526	24.5173%	789,593	24.1479%
100%	5,036,364	100.0000%	5,710,214	100.0000%	7,378,782	100.0000%	4,374,567	100.0000%	3,269,814	100.0000%

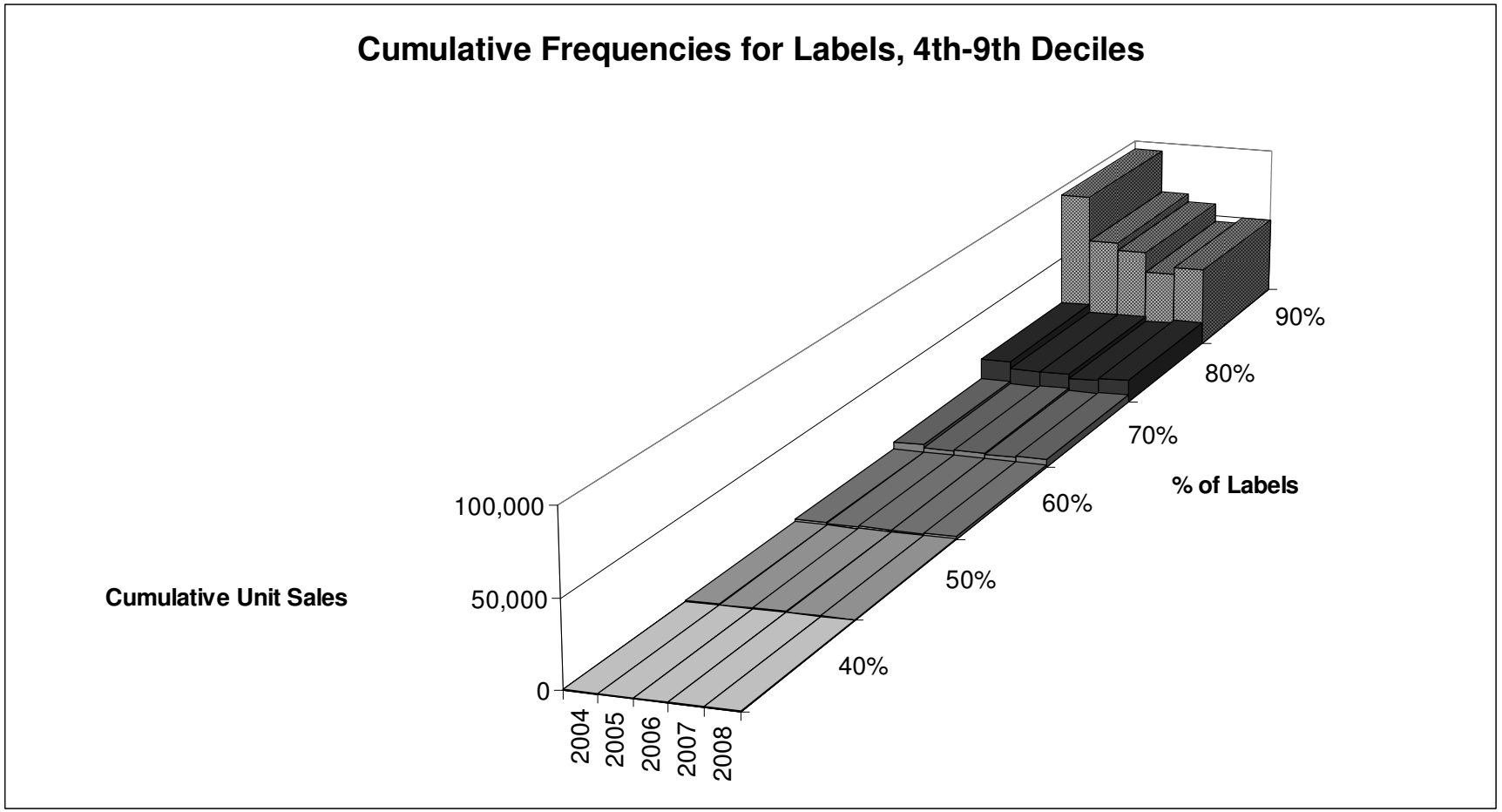

Figure 26: Cumulative Frequencies for 4th through 9th Deciles of Label Sales, 2004-2008.

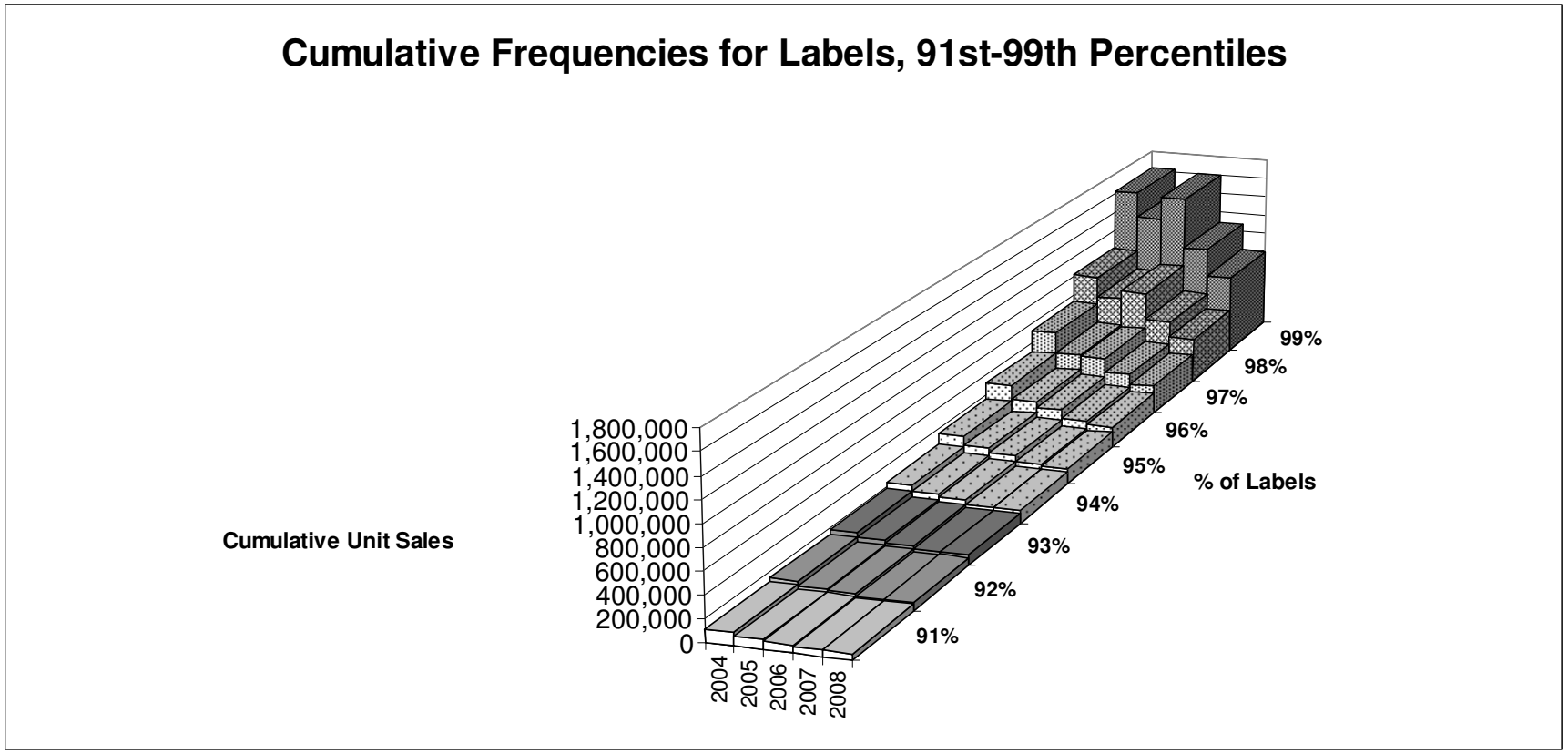

Figure 27: Cumulative Frequencies for 91st through 99th Percentiles of Label Unit Sales, 2004-2008.

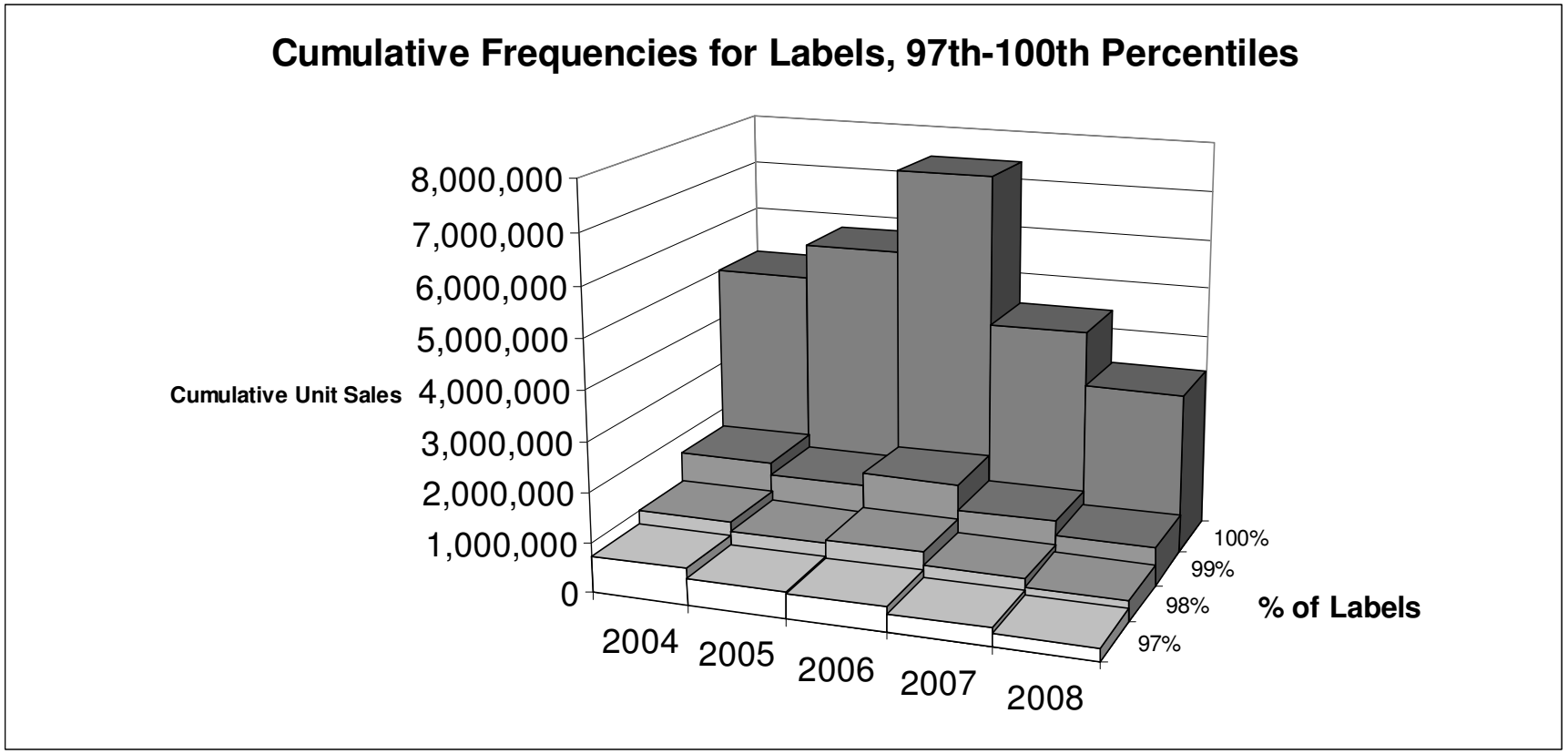

Figure 28: Cumulative Frequencies for 97th through 100th Percentiles of Label Unit Sales, 2004-2008.

Another way to compare relative shifts in industry concentration over time is with Gini coefficients. A Gini coefficient has values between zero and one and measures the degree of consolidation of a variable across individuals contributing to it. It compares the Lorenz curve of an ordinally-ranked distribution with the line of perfect equality, which is the 45 degree line. Specifically, the area in between the Lorenz curve and the 45 degree line is divided by the entire area under the 45 degree line and Lorenz curve combined. The Lorenz curve is a graph of the cumulative share or distribution function of rank-ordered observations. A higher Gini coefficient implies a higher concentration of the distribution. For example, a Gini coefficient of unity implies one individual sells 100% of the total sales while the remainder of individuals sell 0%. A lower Gini coefficient implies a lower concentration in the distribution of sales. For example, a Gini coefficient of 0 implies that each individual has equal sales, each individual contributes equally to total sales, and the Lorenz curve is superimposed upon the 45 degree line. The Gini coefficient results are displayed in Figure 17 for artists, Figure 18 for titles, and Figure 19 for labels. All Gini coefficients are in the upper nineties which implies an extremely high concentration of sales. As with the analysis above, there is a trend of rising inequality and then declining inequality across the five year time span. But for Gini coefficients, the peak concentration years range from 2005 to 2007. Many of these year-to-year changes for this measure are quite small which throws doubt upon their economic significance.

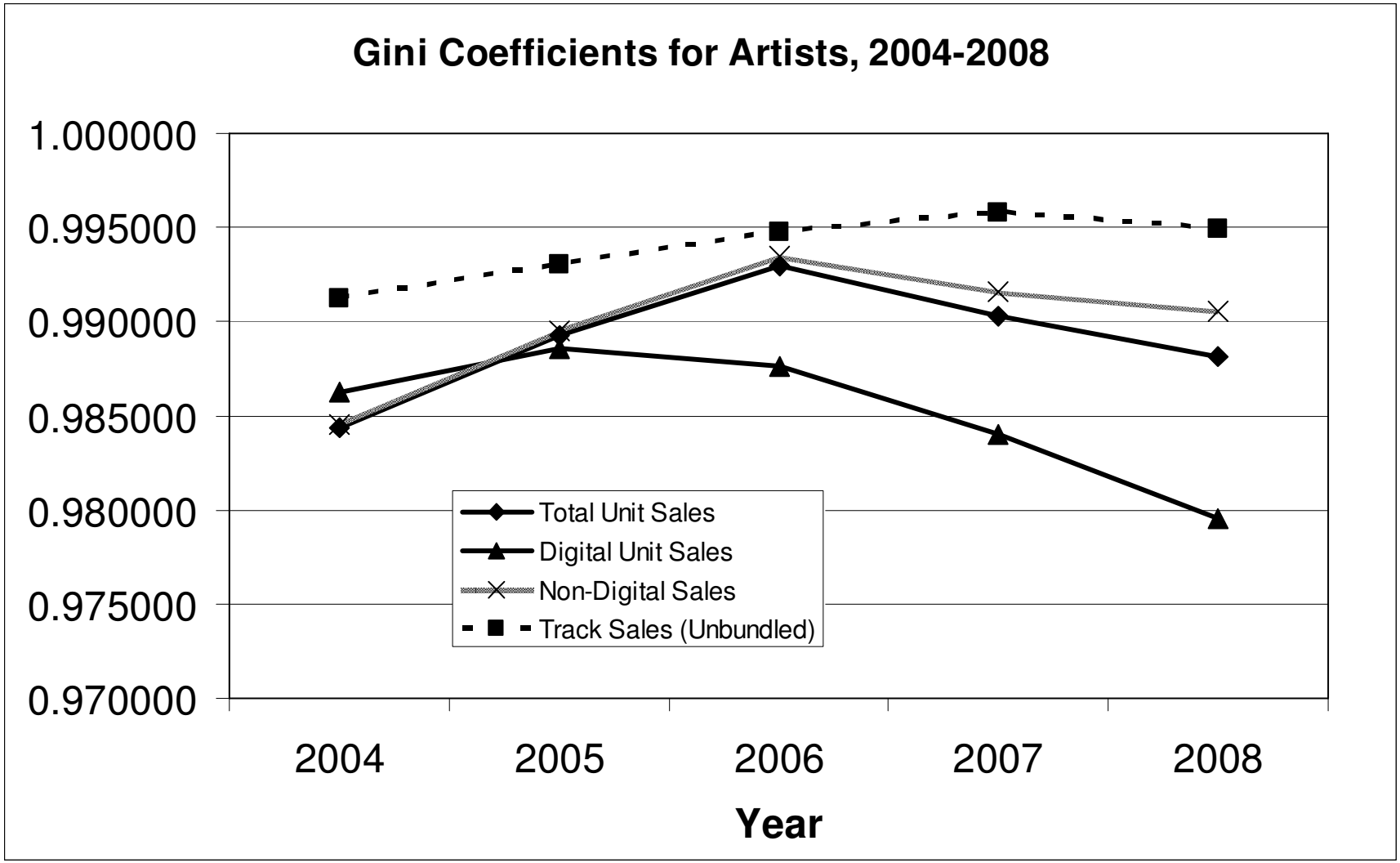

Figure 29: Gini Coefficients for Artists, 2004-2008.
From 2004 through 2008 for total unit sales, bundled digital unit sales, non-digital unit sales, and unbundled track sales.

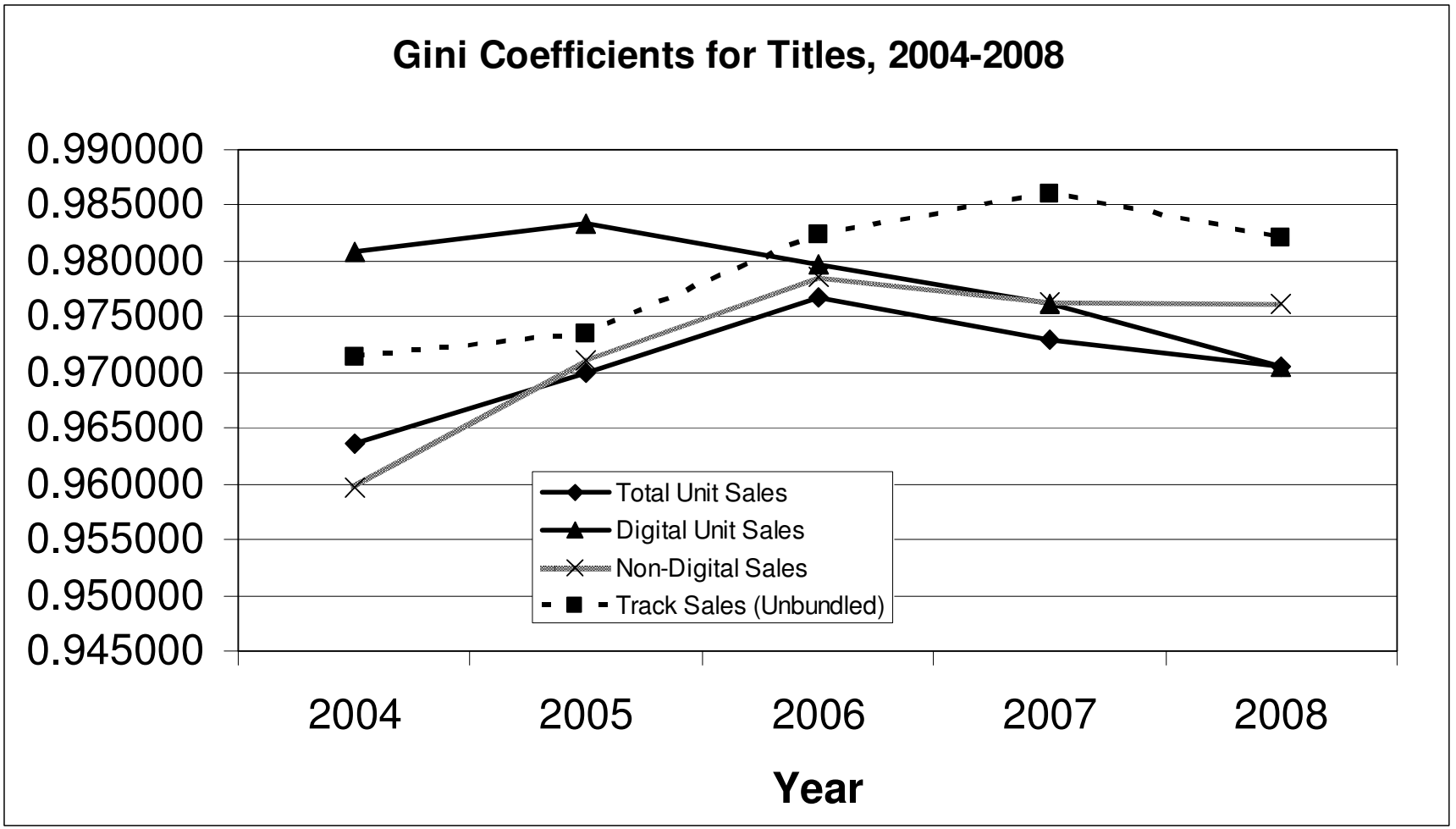

Figure 30: Gini Coefficients for Titles, 2004-2008.
From 2004 through 2008 for total unit sales, bundled digital unit sales, non-digital unit sales, and unbundled track sales.

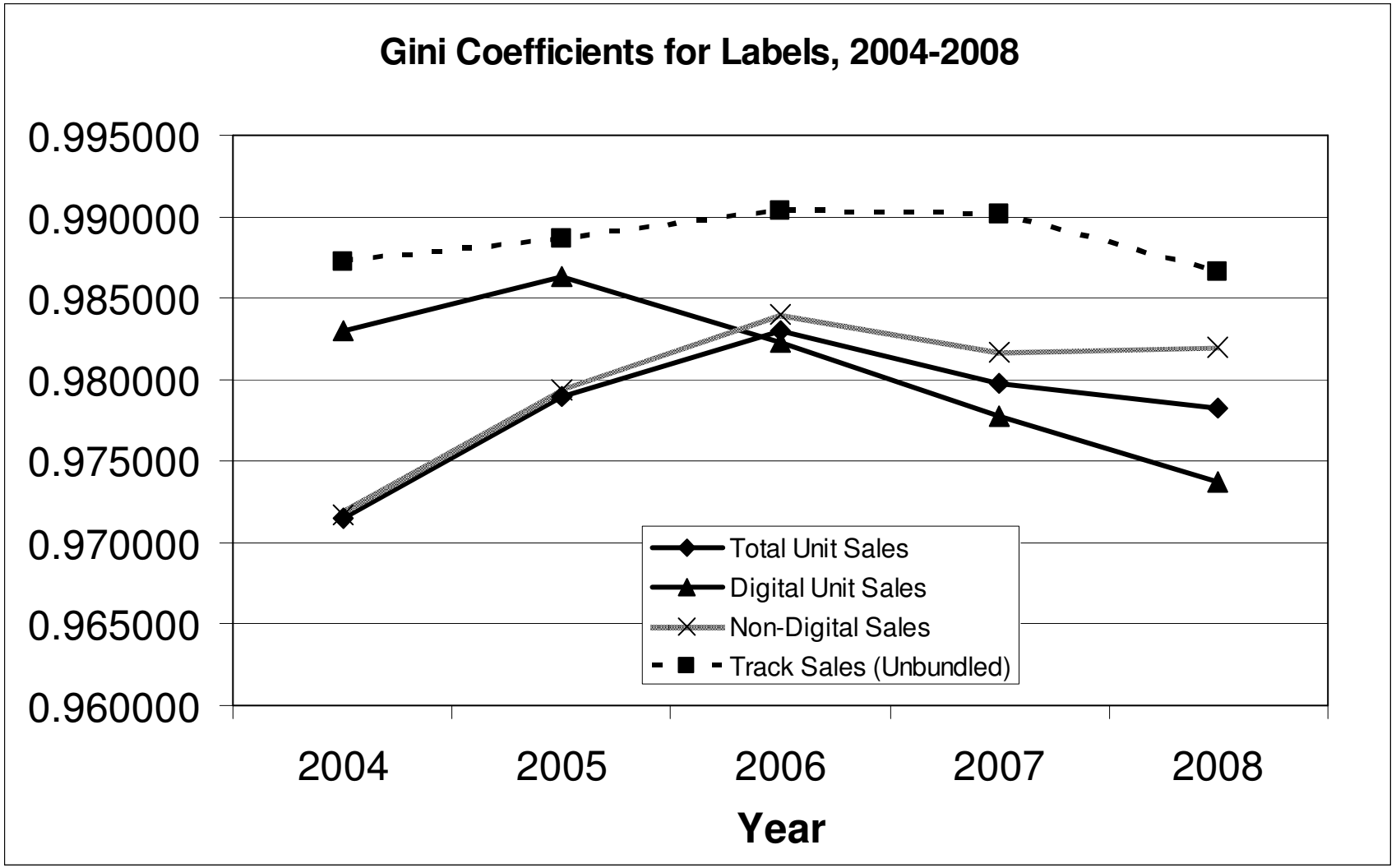

Figure 31: Gini Coefficients for Labels, 2004-2008.
From 2004 through 2008 for total unit sales, bundled digital unit sales, non-digital unit sales, and unbundled track sales.

Figures 20, 21, and 22 show the same Gini coefficient calculations but with zero sales dropped. A similar pattern is observed, however the decreasing concentration observed above in the later years is much less pronounced. In all three figures, there is rising inequality and then a leveling out in which consolidation of sales remains about the same or declines only slightly in the final year or two. So when only non-zero sales are taken into account, there is a superstar effect for the first three years – 2004, 2005, and 2006 – for albums, titles, and labels. Then, in 2007 and 2008, there is a slight decline in concentration. All of the Gini coefficients for artists, titles, and labels end up higher at the end of the five year period. This is a clear superstar effect, regardless of the slight decline in concentration observed in the final two years.

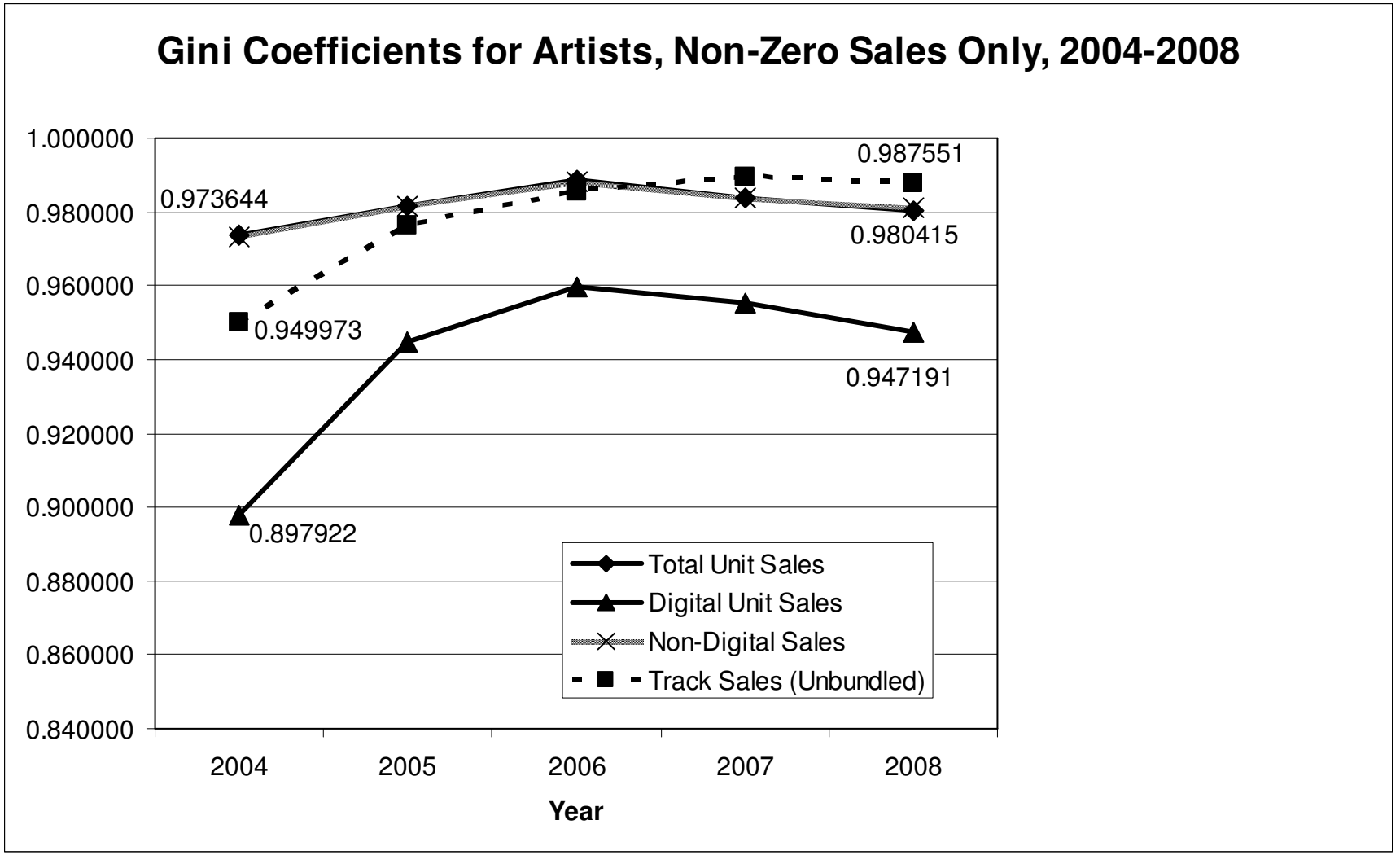

Figure 32: Gini Coefficients for Artists, Non-zero Sales Only, 2004-2008.
From 2004 through 2008 for only *non-zero* total unit sales, bundled digital unit sales, non-digital unit sales, and unbundled track sales.

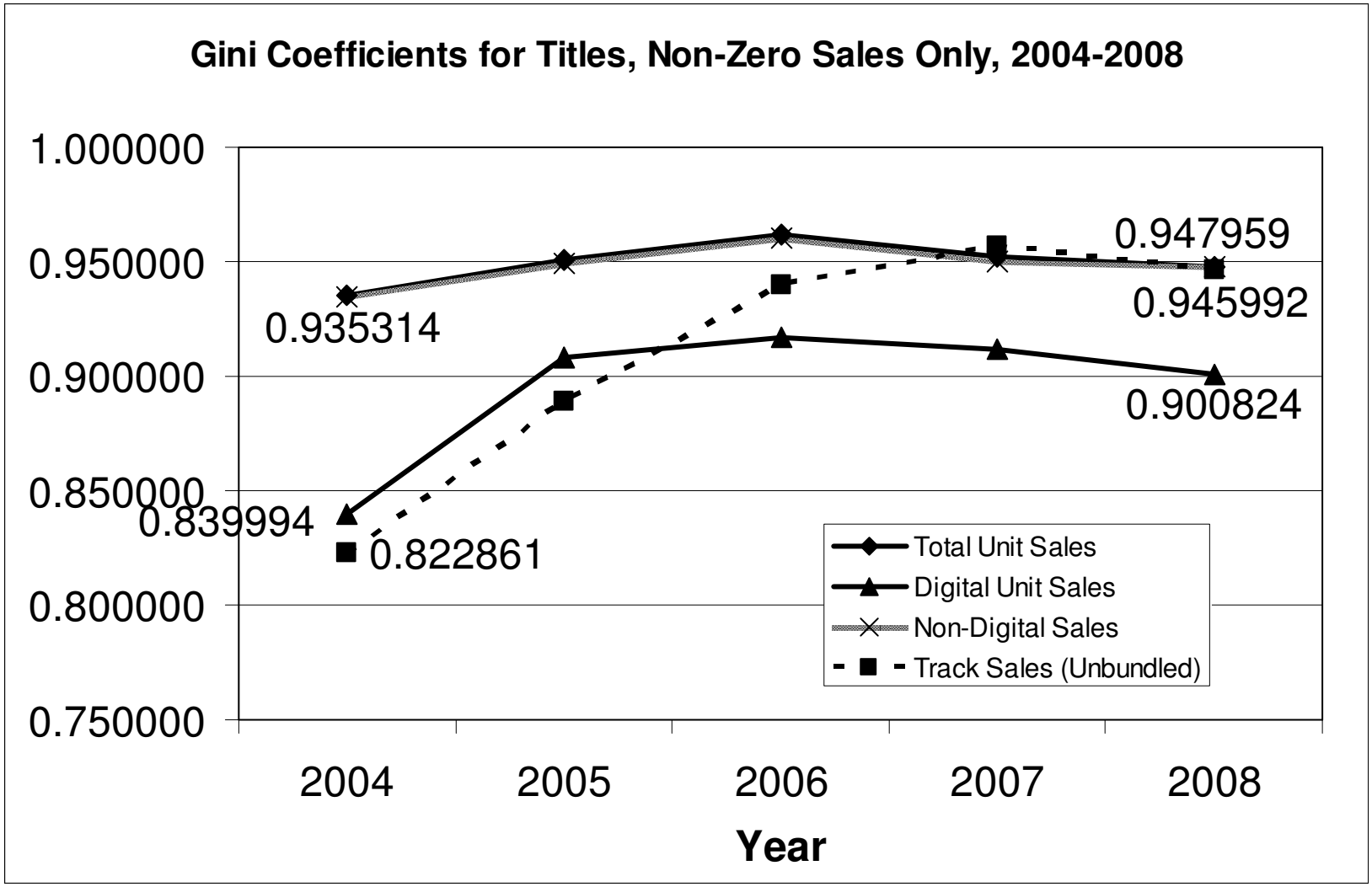

Figure 33: Gini Coefficients for Titles, Non-zero Sales Only, 2004-2008.
From 2004 through 2008 for only *non-zero* total unit sales, bundled digital unit sales, non-digital unit sales, and unbundled track sales.

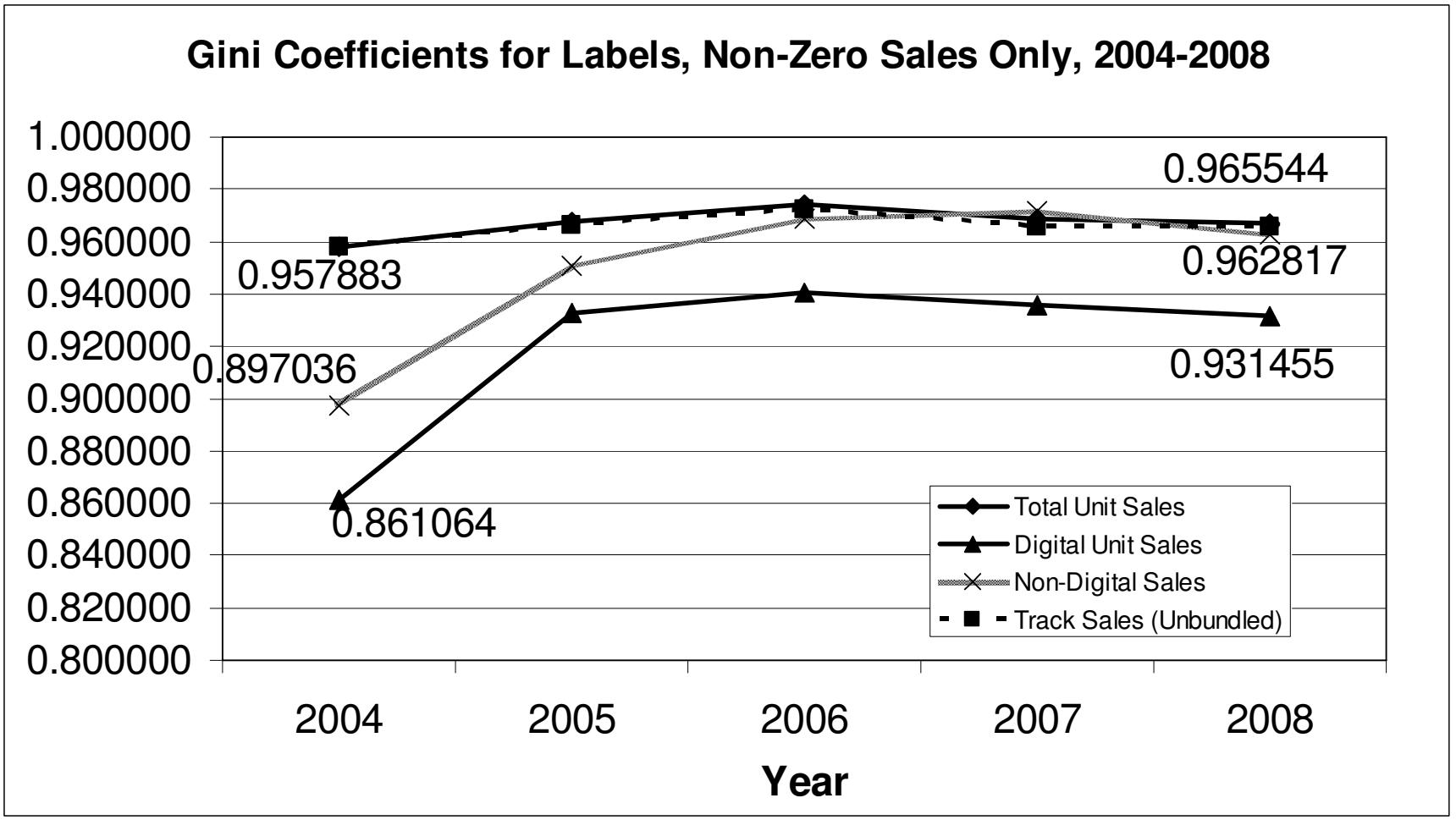

Figure 34: Gini Coefficients for Labels, Non-zero Sales Only, 2004-2008.
From 2004 through 2008 for only *non-zero* total unit sales, bundled digital unit sales, non-digital unit sales, and unbundled track sales.

B. Absolute Sales Changes

This section looks at absolute changes by analyzing the location, scale, skewness, kurtosis, and inter-quartile measures from year to year. This illuminates the distributional shapes and, for example, how tails shift relative to the median. In this section, quantile is abbreviated with a Q, such that $Q_{.75}$ is the 75th percentile or 3rd quartile. $Q_{.50}$ stands for the median or 2nd quartile. Table 7 shows absolute distributional changes from year to year across artists. For albums, singles, and videos together, the number of artists with zero yearly sales increases every year. For all formats, this number trends up from 477 artists in 2004 to 806 artists by 2008. The percentage of artists who sell one or more units, i.e. the percentage of artists that have non-zero sales, remains fairly constant at around 60%. For all formats, the number of artists reaching sales levels above the 50th percentile of unit sales increases steadily from 573 artists in 2004 to 947 in 2008. The number of artists with sales levels above the 90th percentile rises consistently from 117 in 2004 to 204 in 2008. The number of artists with sales above the 99th percentile rises from 12 to 21 over this five year period.

Table 10: Absolute Distributional Changes for Artists: The Distribution Described in Sales Units, 2004-2008.

Across years for all formats. This shows how many sales were equal to zero, greater than zero, greater than the median ($Q_{.50}$), 3rd quartile ($Q_{.75}$), 90th percentile ($Q_{.90}$), 99th percentile ($Q_{.99}$) numbers of sales, as well as the percentage of non-zero sales.

		Number of Artists		All Formats						
		N=1172		N=1336		N=1575		N=1795		N=2051
	sales level	2004	sales level	2005	sales level	2006	sales level	2007	sales level	2008
Sales=0		477		546		605		707		806
Sales>0		695		790		970		1088		1245
Q.50	2	573	2	623	2	756	2	822	2	947
Q.75	46	293	31	330	25	391	19	438	16	510
Q.90	717	117	471	134	302	158	210	180	188	204
Q.99	44424	12	32868	14	27513	16	20600	18	16763	21
% non-zero		59.30%		59.13%		61.59%		60.61%		60.70%

Table 8 displays the same absolute changes in the distribution of sales across titles. For all formats together, the number of titles with zero yearly sales increases each year. This number rises from 1,423 titles in 2004 to 3,033 titles in 2008. The percentage of titles which sell one or more units declines slightly. In 2004, over 63% of titles have non-zero sales while in 2008 it is about 57%. Across all formats, the number of titles with sales levels above the 50th percentile increases from 1,922 in 2004 to 2,918 in 2008. The number of titles with sales levels above the 90th percentile nearly doubles from 388 in 2004 to 700 in 2008. The number of titles with sales above the 99th percentile also nearly doubles from 39 in 2004 to 71 in 2008.

Table 11: Absolute Distributional Changes for Titles: The Distribution Described in Sales Units, 2004-2008.

Across years for all formats. This shows how many sales were equal to zero, greater than zero, greater than the median ($Q_{.50}$), 3rd quartile ($Q_{.75}$), 90th percentile ($Q_{.90}$), 99th percentile ($Q_{.99}$) numbers of sales, as well as the percentage of non-zero sales.

		Number of Titles		All Formats						
		N=3882		N=4468		N=5201		N=6088		N=7010
	sales level	2004	sales level	2005	sales level	2006	sales level	2007	sales level	2008
Sales=0		1423		1757		2048		2647		3033
Sales>0		2459		2711		3153		3440		3976
Q.50	4	1922	3	2130	2	2579	1	2987	1	3414
Q.75	75	966	50	1111	41	1292	24	1517	20	1740
Q.90	1004	388	675	447	488	520	305	608	201	700
Q.99	27918	39	20280	45	23915	52	13420	61	8475	71
% non-zero		63.34%		60.68%		60.62%		56.51%		56.73%

Summary statistics for the sales distributions for artists are revealed in Table 9. The location of the median ($Q_{.50}$) and scale, $(Q_{.75}-Q_{.25})/(Q_{.75}+Q_{.25})$, remain the same for each year. Scale remains the same because the 25th percentile is 0. The measure of skewness, $(Q_{.75}+Q_{.25}-2Q_{.5})/(Q_{.75}-Q_{.25})$, is positive and declines, i.e. becomes less skewed, steadily over time from 0.91 in 2004 to 0.75 in 2008. This means that the sales distribution becomes less concentrated which implies the long tail effect. Similarly, kurtosis, $(Q_{.9}-Q_{.1})/(Q_{.75}-Q_{.25})$, generally declines over time from 15.75 in 2004 to 11.06 in 2007 with a slight rise again in 2008 to 11.75. A decline in kurtosis implies a more rounded peak and shorter, thinner tail for the distribution. These numbers provide evidence that the distribution of artist sales becomes less asymmetric and less concentrated. However, the decline in kurtosis means the sales distribution becomes less spread out. The inter-quartile measure $(Q_{.75}-Q_{.25})$, i.e. the difference between the 3rd and 1st quartiles, decreases consistently each year from 45.5 in 2004 to 16 in 2008. This implies that the middle class of artists is shrinking, which supports the hybrid hypothesis. The left tail inter-quartile measure $(Q_{.50}-Q_{.25})/Q_{.5}$ remains constant at 1. This is consistent with a long, flat tail. The right tail inter-quartile measure $(Q_{.75}-Q_{.50})/Q_{.50}$ declines each year starting from 21.75 in 2004 and ending at 7 in 2008. This also implies that the middle class of artists is shrinking, which supports the hybrid hypothesis.

Table 12: Summary Statistics for Artists: Location, Scale, Skewness, Kurtosis, Inter-quartile, Left Tail, Right Tail, 2004-2008.

This displays the location of the median, scale, skewness, kurtosis, and inter-quartile measures for each year.

Distributions for Artists		Sales Levels				
All Formats		N=1172	N=1336	N=1575	N=1795	N=2051
Measure	Definition	2004	2005	2006	2007	2008
Location	Q.5	2.00	2.00	2.00	2.00	2.00
Scale	(Q.75-Q.25)/(Q.75+Q.25)	1.00	1.00	1.00	1.00	1.00
Skewness	(Q.75+Q.25-2Q.5)/(Q.75-Q.25)	0.91	0.87	0.84	0.79	0.75
Kurtosis	(Q.9-Q.1)/(Q.75-Q.25)	15.75	15.19	12.33	11.06	11.75
Inter-quartile	(Q.75-Q.25)	45.50	31.00	24.50	19.00	16.00
Left Tail	(Q.5-Q.25)/Q.5	1.00	1.00	1.00	1.00	1.00
Right Tail	(Q.75-Q.5)/Q.5	21.75	14.50	11.25	8.50	7.00

Table 10 lists summary statistics for the sales distributions for titles. The location of the median, ($Q_{.50}$), declines from 4 in 2004 to 1 in 2007. Scale, $(Q_{.75}-Q_{.25})/(Q_{.75}+Q_{.25})$, remains the same for each year. Scale remains the same because the 25th percentile is 0. Skewness, $(Q_{.75}+Q_{.25}-2Q_{.5})/(Q_{.75}-Q_{.25})$, is positive and basically remains the same at around 0.9. Kurtosis, (Q.9-Q.1)/(Q.75-Q.25), does not show any obvious trend. The inter-quartile measure ($Q_{.75}-Q_{.25}$) decreases consistently each year from 75 in 2004 to 20 in 2008. This supports the hybrid hypothesis because sales are declining for the middle class. The left tail inter-quartile measure remains constant at 1. This suggests a long, flat tail. The right tail inter-quartile measure does not show any obvious trend.

Table 13: Summary Statistics for Titles: Location, Scale, Skewness, Kurtosis, Inter-quartile, Left Tail, Right Tail, 2004-2008.

This displays the location of the median, scale, skewness, kurtosis, and inter-quartile measures for each year.

Distributions for Titles		Sales Levels				
All Formats		N=3882	N=4468	N=5201	N=6088	N=7010
Measure	Definition	2004	2005	2006	2007	2008
Location	Q.5	4.00	3.00	2.00	1.00	1.00
Scale	(Q.75-Q.25)/(Q.75+Q.25)	1.00	1.00	1.00	1.00	1.00
Skewness	(Q.75+Q.25-2Q.5)/(Q.75-Q.25)	0.89	0.88	0.90	0.92	0.90
Kurtosis	(Q.9-Q.1)/(Q.75-Q.25)	13.38	13.50	11.91	12.69	10.05
Inter-quartile	(Q.75-Q.25)	75.00	50.00	41.00	24.00	20.00
Left Tail	(Q.5-Q.25)/Q.5	1.00	1.00	1.00	1.00	1.00
Right Tail	(Q.75-Q.5)/Q.5	17.75	15.67	19.50	23.00	19.00

C. Regression Analysis

There is much evidence that the sales distributions are changing from year to year. The next step is to test the statistical significance of these differences. In this section, unit sales are regressed on sales rank in the following log-linear form for annual sales:

$$\ln(\text{Sales}+1)= \beta_0+ \beta_1 \cdot \ln(\text{Sales Rank})+\varepsilon$$

'Sales' are the total unit sales for each year from 2004 to 2008. 'Sales Rank' represents artists or titles ordinally ranked by the sum of their yearly unit sales. Sales is regressed on sales rank for artists as well as titles for total unit sales, digital unit sales, non-digital, and unbundled track sales. The coefficients on Sales Rank are then compared between years. Subsequently, the same regressions are performed with zero sales dropped. Firstly, β_0 can be considered a measure of overall sales in a given year. Secondly, β_1 can be considered a magnitude of how rapidly the share of yearly sales drops off as sales rank increases. The long tail hypothesis implies that β_1 will decrease in absolute value over time due to the institutional and technological changes as consumers diversify consumption as more variety is made available to them. So, the slope with which sales drops off will be less steep or become less negative over time. The intuition is that less popular artists will gain a larger share of sales over time. It is the same analogy for regressions with titles instead of artists.

The superstar hypothesis would be just the opposite effect, with an absolute value of β_1 increasing over time. The slope will become steeper or more negative from year to year.

The same institutional and technological changes on the supply side allows consumers not only to 'ride down the long tail' to seek out unique, niche, and less popular content but also buy more of what the superstars have available because superstars can reach larger audiences than they could before the institutional constraints and bottlenecks were lifted. Tables 11, 12, 13, and 14 describe the data to be regressed aggregated across artists and titles from 2004 to 2008.

Because both variables are in log form, β_1 represents an elasticity. These elasticities will be negative because as sales rank increases for an artist, they are moving down the list with lower sales. As rank increases, sales decline. The β_1 then can be referred to as a rank elasticity of sales or the elasticity of sales with respect to rank.

Table 11 lists the number of observations, mean, standard deviation, minimum and maximum for sales aggregated across artists. The number of artists captured in the sample rises from 1,172 in 2004 to 2,051 in 2008. The formats are broken into total unit sales, digital unit sales, non-digital sales, and unbundled track sales.

Table 14: Synopsis of Artist Sales: Number of Observations, Mean, Standard Deviation, Minimum, Maximum.

Variable	Obs	Mean	Std. Dev.	Min	Max
YTD Total Unit Sales 2008	2051	1594.255	28492.13	0	1182376
YTD Digital Unit Sales 2008	2051	186.256	2038.171	0	62122
YTD Non-Digital Unit Sales 2008	2051	1407.999	26726.05	0	1120254
YTD Track Sales 2008	2051	20936.99	382450	0	1.36e+07
YTD Total Unit Sales 2007	1795	2437.085	46239.81	0	1822934
YTD Digital Unit Sales 2007	1795	200.0401	2375.612	0	68663
YTD Non-Digital Unit Sales 2007	1795	2237.045	44158.74	0	1754271
YTD Track Sales 2007	1795	22250.67	490889.2	0	1.93e+07
YTD Total Unit Sales 2006	1575	4684.949	129048.2	0	5092834
YTD Digital Unit Sales 2006	1575	174.7733	3101.211	0	113096
YTD Non-Digital Unit Sales 2006	1575	4510.175	126115.8	0	4979738
YTD Track Sales 2006	1575	13547.8	241802.3	0	7259474
YTD Total Unit Sales 2005	1336	4274.117	80487.39	0	2826446
YTD Digital Unit Sales 2005	1336	113.9401	1789.257	-1	48433
YTD Non-Digital Unit Sales 2005	1336	4160.177	79134.11	0	2785806
YTD Track Sales 2005	1336	6286.979	103021.9	0	3175213
YTD Total Unit Sales 2004	1172	4297.239	58436.92	0	1766603
YTD Digital Unit Sales 2004	1172	36.65614	424.3298	0	11877
YTD Non-Digital Unit Sales 2004	1172	4260.583	58027.55	0	1754726
YTD Track Sales 2004	1172	2505.443	31339.8	0	760239

Table 12 lists the number of observations, mean, standard deviation, minimum and maximum for sales aggregated across artists, but with artists who sell zero dropped. The number of artists captured in the sample with non-zero sales is variable.

Table 15: Synopsis of Non-zero Artist Sales: Number of Observations, Mean, Standard Deviation, Minimum, Maximum.

Variable	Obs	Mean	Std. Dev.	Min	Max
YTD Total Unit Sales 2008	1245	2626.359	36538.52	1	1182376
YTD Digital Unit Sales 2008	793	481.7289	3257.299	1	62122
YTD Non-Digital Unit Sales 2008	1020	2831.182	37854.22	1	1120254
YTD Track Sales 2008	831	51674.81	599728.4	1	1.36e+07
YTD Total Unit Sales 2007	1088	4020.742	59349.89	1	1822934
YTD Digital Unit Sales 2007	643	558.4323	3945.857	1	68663
YTD Non-Digital Unit Sales 2007	927	4331.71	61390.29	1	1754271
YTD Track Sales 2007	729	54787.31	769440.9	1	1.93e+07
YTD Total Unit Sales 2006	970	7607.004	164404.5	1	5092834
YTD Digital Unit Sales 2006	484	568.7355	5578.263	1	113096
YTD Non-Digital Unit Sales 2006	872	8146.245	169448.9	1	4979738
YTD Track Sales 2006	578	36916.57	398286.1	1	7259474
YTD Total Unit Sales 2005	790	7228.127	104593.8	1	2826446
YTD Digital Unit Sales 2005	277	549.5487	3904.481	1	48433
YTD Non-Digital Unit Sales 2005	752	7390.952	105394.4	1	2785806
YTD Track Sales 2005	390	21536.93	189986.8	1	3175213
YTD Total Unit Sales 2004	695	7246.567	75766.62	1	1766603
YTD Digital Unit Sales 2004	158	271.9051	1130.748	1	11877
YTD Non-Digital Unit Sales 2004	682	7321.705	75944.39	1	1754726
YTD Track Sales 2004	205	14323.8	73944.65	1	760239

Table 13 lists the number of observations, mean, standard deviation, minimum and maximum for sales aggregated across titles. The number of individual titles included in the sample rises from 3,877 in 2004 to 7,010 in 2008. Once again, the formats are broken into total unit sales, digital unit sales, non-digital sales, and unbundled track sales.

Table 16: Synopsis of Title Sales: Number of Observations, Mean, Standard Deviation, Minimum, Maximum.

Variable	Obs	Mean	Std. Dev.	Min	Max
YTD Total Unit Sales 2008	7010	466.4504	4968.971	-9	233615
YTD Digital Unit Sales 2008	7010	54.49515	520.1796	0	15738
YTD Non-Digital Unit Sales 2008	7010	411.9552	4654.102	-9	222817
YTD Track Sales 2008	7010	6125.788	61534.43	0	1955003
YTD Total Unit Sales 2007	6087	718.6737	8572.936	-1	411073
YTD Digital Unit Sales 2007	6087	58.98998	668.4988	0	31223
YTD Non-Digital Unit Sales 2007	6087	659.5754	8079.362	-3	395100
YTD Track Sales 2007	6087	6561.516	113033.3	0	6885512
YTD Total Unit Sales 2006	5199	1419.272	23855.05	0	1430885
YTD Digital Unit Sales 2006	5199	52.94634	721.0653	0	35509
YTD Non-Digital Unit Sales 2006	5199	1365.8	23518.49	0	1429360
YTD Track Sales 2006	5199	4104.208	55564.72	0	2453507
YTD Total Unit Sales 2005	4465	1278.885	15969.11	0	778298
YTD Digital Unit Sales 2005	4465	34.09272	604.1053	-1	34429
YTD Non-Digital Unit Sales 2005	4465	1243.956	15615.6	0	773583
YTD Track Sales 2005	4465	1881.166	13149.16	0	313567
YTD Total Unit Sales 2004	3877	1299.036	11058.91	0	380157
YTD Digital Unit Sales 2004	3877	11.08099	119.3603	0	3788
YTD Non-Digital Unit Sales 2004	3877	1286.296	10955.81	0	376369
YTD Track Sales 2004	3877	757.3843	4934.573	0	82801

Table 14 lists the number of observations, mean, standard deviation, minimum and maximum for sales aggregated across titles, but with titles which sell zero dropped. As with

Table 12 for artists, the number of titles captured in the sample with non-zero sales is variable.

Table 17: Synopsis of Non-Zero Title Sales: Number of Observations, Mean, Standard Deviation, Minimum, Maximum.

Variable	Obs	Mean	Std. Dev.	Min	Max
YTD Total Unit Sales 2008	3976	822.3908	6575.982	1	233615
YTD Digital Unit Sales 2008	2085	183.2187	941.5166	1	15738
YTD Non-Digital Unit Sales 2008	3186	906.4077	6871.59	1	222817
YTD Track Sales 2008	2319	18517.37	105923.3	1	1955003
YTD Total Unit Sales 2007	3439	1272.047	11375.34	1	411073
YTD Digital Unit Sales 2007	1640	218.9463	1274.507	1	31223
YTD Non-Digital Unit Sales 2007	2902	1383.7	11660.3	1	395100
YTD Track Sales 2007	1956	20419.2	198722.9	1	6885512
YTD Total Unit Sales 2006	3151	2341.731	30608.56	1	1430885
YTD Digital Unit Sales 2006	1263	217.9477	1451.051	1	35509
YTD Non-Digital Unit Sales 2006	2804	2533.354	31986.93	1	1429360
YTD Track Sales 2006	1513	14102.96	102337.6	1	2453507
YTD Total Unit Sales 2005	2708	2108.648	20464.12	1	778298
YTD Digital Unit Sales 2005	806	188.8648	1412.256	1	34429
YTD Non-Digital Unit Sales 2005	2537	2190.775	20674.76	1	773583
YTD Track Sales 2005	1071	7842.581	25971.86	1	313567
YTD Total Unit Sales 2004	2454	2052.308	13845.56	1	380157
YTD Digital Unit Sales 2004	466	92.19099	333.558	1	3788
YTD Non-Digital Unit Sales 2004	2395	2084.928	13889.52	1	376369
YTD Track Sales 2004	625	4698.206	11519.83	1	82801

Regression results for artists are displayed in Table 15. The β_1s for each year are then graphed for ease of comparison in Figure 23. To interpret these coefficients economically, recall that they represent elasticities. If β_1 increases over time – decreases in absolute value – then sales become less elastic with respect to rank over time. For example, a 1% increase (decrease) in sales rank causes a decline (rise) in the average artist's sales of 2.83% in 2004, and a decline (rise) in sales of 2.34% in 2008. All coefficients are statistically different from zero, $p<0.01$, and all R^2s are high. β_1 increases, or decreases in absolute value, for artists from 2004 through 2008 for all unit sales and non-digital sales. This implies a long tail effect: less concentration of sales over time. However, the β_1 for digital units and for tracks declines, or increases in absolute value, each year from 2004 to 2008. This implies a superstar effect: less concentration of sales over time for these particular formats.

Table 18: Regression Results for Artists, 2004-2008.

All unit sales, bundled digital sales, unbundled track sales, and non-digital sales. One asterisk (*) indicates significance at 5% and two (**) significance at 1%.

Artists, All Unit Sales

	2004	2005	2006	2007	2008
β_0	19.477740	18.762840	18.168920	17.529170	17.399350
Std. Err.	0.094356	0.076053	0.057714	0.047481	0.044571
β_1	-2.831656**	-2.683175**	-2.531821**	-2.406398**	-2.344046**
Std. Err.	0.015343	0.012112	0.008960	0.007226	0.006650
R^2	0.966800	0.973500	0.980700	0.984100	0.983800

Artists, Bundled Digital Sales

	2004	2005	2006	2007	2008
β_0	6.972735	9.433881	11.332410	12.480830	13.344460
Std. Err.	0.125508	0.115027	0.087198	0.075980	0.068072
β_1	-1.082853**	-1.421365**	-1.646286**	-1.769906**	-1.848238**
Std. Err.	0.020408	0.018321	0.013537	0.011564	0.010157
R^2	0.706400	0.818700	0.903900	0.928900	0.941700

Artists, Unbundled Track Sales

	2004	2005	2006	2007	2008
β_0	12.862030	16.351370	18.267020	19.259580	19.811540
Std. Err.	0.185478	0.137456	0.108006	0.092543	0.090689
β_1	-1.981948**	-2.429178**	-2.622732**	-2.700930**	-2.730577**
Std. Err.	0.030159	0.021891	0.016768	0.014085	0.013532
R^2	0.786800	0.902300	0.939600	0.953500	0.952100

Artists, Non-Digital Sales

	2004	2005	2006	2007	2008
β_0	19.42923	18.58049	17.79645	16.78220	16.20026
Std. Err.	0.09547	0.07770	0.06307	0.05824	0.05547
β_1	-2.82986**	-2.66738**	-2.50383**	-2.33471**	-2.21849**
Std. Err.	0.01552	0.01237	0.00979	0.00886	0.00828
R^2	0.96600	0.97210	0.97650	0.97480	0.97230

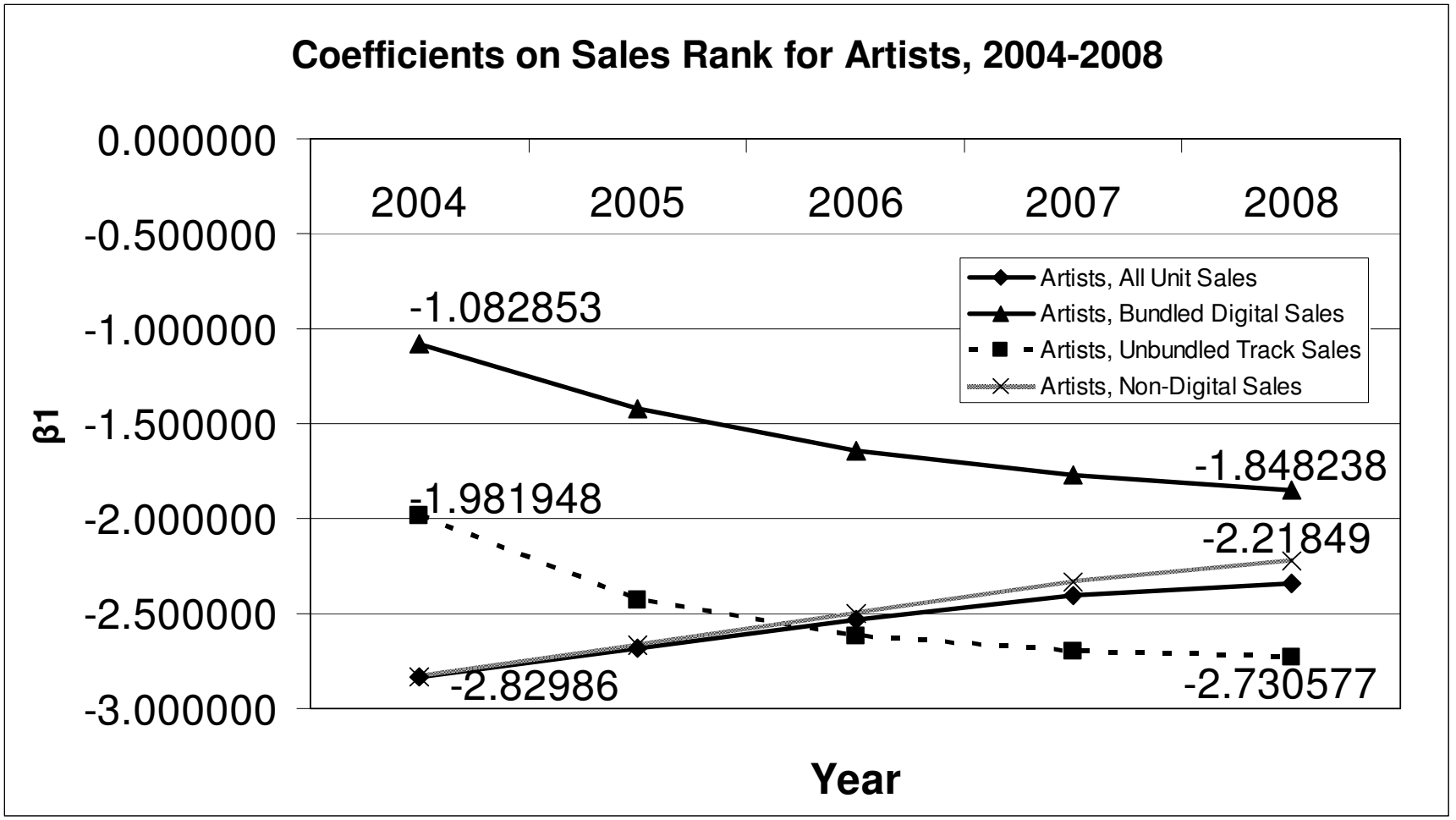

Figure 35: Coefficients on Sales Rank for Artists, 2004-2008.
For all unit sales, bundled digital albums and singles, unbundled digital track sales, and non-digital unit sales for each year.

For artists, the tests for statistically significant differences between the β_1s between each pair of consecutive years are shown in Table 15. These are all highly significant, $p<0.01$. So, the differences in concentration implied by the coefficients on sales rank each year are significant.

Table 19: Test of Significance Between Years for Artists.

Results for the test of significance for the coefficient on sales rank (β_1) between each pair of consecutive years for artists.

Artists, All Unit Sales

	2005-2004	2006-2005	2007-2006	2008-2007
chi2(1)	130.28	157.62	172.47	168.57
Prob > chi2	0.00000	0.00000	0.00000	0.00000

Artists, Bundled Digital Sales

	2005-2004	2006-2005	2007-2006	2008-2007
chi2(1)	246.43	156.42	168.33	70.48
Prob > chi2	0.00000	0.00000	0.00000	0.00000

Artists, Unbundled Track Sales

	2005-2004	2006-2005	2007-2006	2008-2007
chi2(1)	154.95	142.11	84.81	25.11
Prob > chi2	0.00000	0.00000	0.00000	0.00000

Artists, Non-Digital Sales

	2005-2004	2006-2005	2007-2006	2008-2007
chi2(1)	137.16	165.77	221.24	505.85
Prob > chi2	0.00000	0.00000	0.00000	0.00000

Regression results for titles are displayed in Table 16. The β_1s for each year are shown in Figure 24. All coefficients are statistically different from zero, $p<0.01$, and all R^2s are high. β_1 increases, or decreases in absolute value, for titles from 2004 through 2008 for all unit sales and non-digital sales. This implies a long tail effect: less concentration of sales over time. However, the β_1 on digital units and tracks declines, or increases in absolute value each year. This implies a superstar effect: more concentration of sales over time for these formats aggregated by title. This mirrors the pattern observed for artists.

Table 20: Regression Results for Titles, 2004-2008.

All unit sales, bundled digital sales, unbundled track sales, and non-digital sales. One asterisk (*) indicates significance at 5% and two (**) significance at 1%.

Titles, All Unit Sales

	2004	2005	2006	2007	2008
β_0	22.340320	21.841250	21.688680	20.493980	19.774060
Std. Err.	0.083884	0.070414	0.058319	0.048897	0.042992
β_1	-2.727106**	-2.637747**	-2.577229**	-2.408078**	-2.284445**
Std. Err.	0.011441	0.009424	0.007651	0.006286	0.005429
R^2	0.936200	0.946100	0.956200	0.960200	0.961900

Titles, Bundled Digital Sales

	2004	2005	2006	2007	2008
β_0	6.936466	9.465947	11.487680	12.515120	13.183700
Std. Err.	0.069629	0.067104	0.059696	0.054345	0.049074
β_1	-0.908565**	-1.208017**	-1.427520**	-1.520678**	-1.569924**
Std. Err.	0.009497	0.008981	0.007832	0.006986	0.006197
R^2	0.702600	0.802200	0.864700	0.886200	0.901600

Titles, Unbundled Track Sales

	2004	2005	2006	2007	2008
β_0	16.361080	20.106040	22.145970	23.380800	24.233630
Std. Err.	0.140612	0.121996	0.104764	0.092178	0.087964
β_1	-2.122573**	-2.538734**	-2.728928**	-2.817669**	-2.869431**
Std. Err.	0.019178	0.016328	0.013744	0.011850	0.011108
R^2	0.759700	0.844200	0.883500	0.902800	0.905000

Titles, Non-Digital Sales

	2004	2005	2006	2007	2008
β_0	22.33288	21.71379	21.36902	19.71512	18.62412
Std. Err.	0.08396	0.07054	0.05935	0.05321	0.04711
β_1	-2.73357**	-2.63770**	-2.56475**	-2.34554**	-2.18466**
Std. Err.	0.01145	0.00944	0.00779	0.00684	0.00595
R^2	0.93630	0.94590	0.95430	0.95080	0.95060

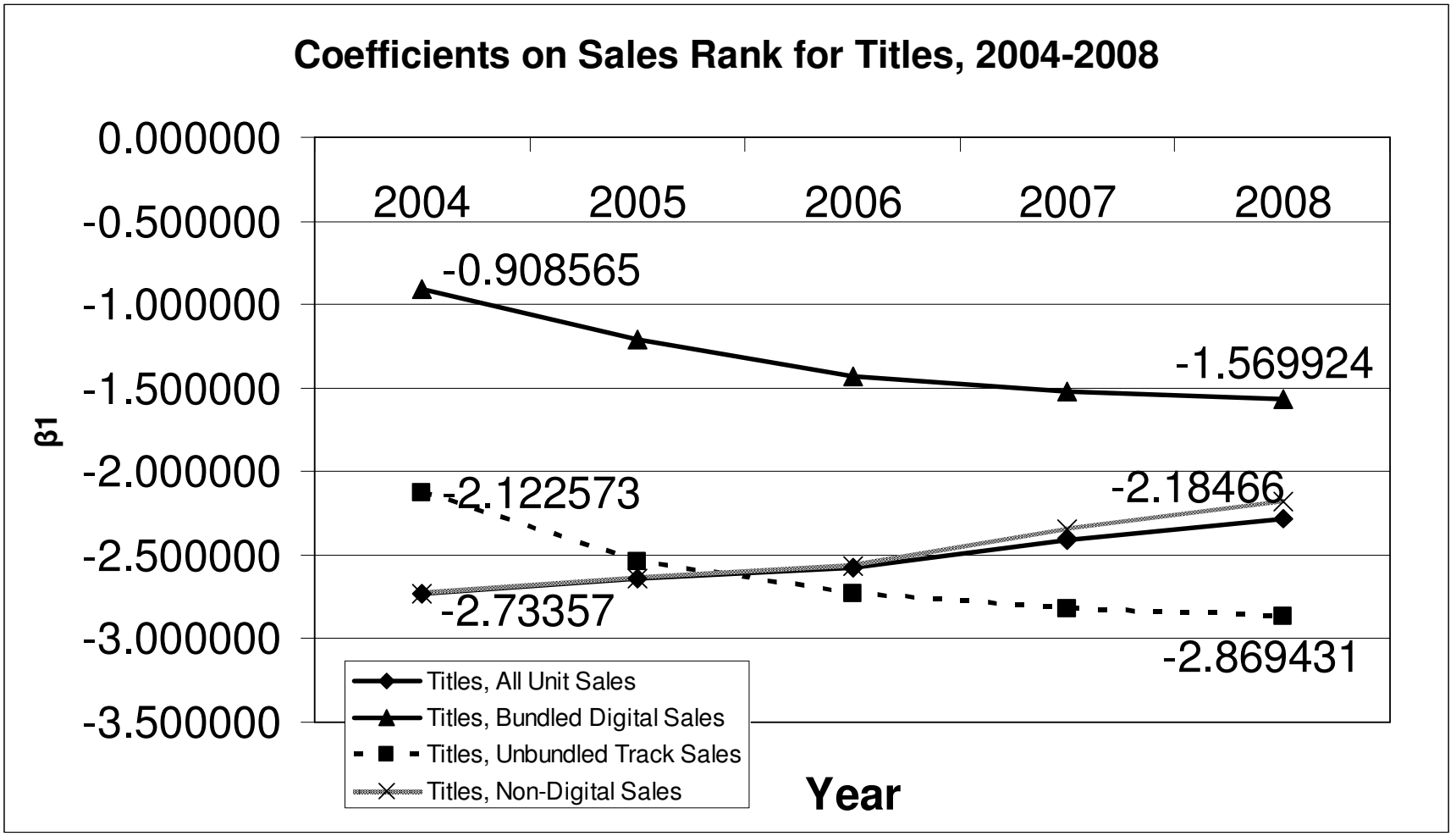

Figure 36: Coefficients on Sales Rank for Titles, 2004-2008.
For all unit sales, bundled digital albums and singles, unbundled digital track sales, and non-digital unit sales for each year.

For titles, the test for statistically significant differences between the β_1s between each pair of consecutive years are shown in Table 17. These are all highly significant, $p<0.01$. So, the differences in concentration implied by the coefficients on sales rank each year are significant.

Table 21: Test of Significance Between Years for Titles.

Results for the test of significance for the coefficient on sales rank (β_1) between each pair of consecutive years for titles.

Titles, All Unit Sales

	2005-2004	2006-2005	2007-2006	2008-2007
chi2(1)	118.66	99.19	1033.95	992.25
Prob > chi2	0.00000	0.00000	0.00000	0.00000

Titles, Bundled Digital Sales

	2005-2004	2006-2005	2007-2006	2008-2007
chi2(1)	1034.05	2744.91	3945.67	155.65
Prob > chi2	0.00000	0.00000	0.00000	0.00000

Titles, Unbundled Track Sales

	2005-2004	2006-2005	2007-2006	2008-2007
chi2(1)	583.05	431.76	227.27	157.39
Prob > chi2	0.00000	0.00000	0.00000	0.00000

Titles, Non-Digital Sales

	2005-2004	2006-2005	2007-2006	2008-2007
chi2(1)	123.38	118.33	1017.94	1198.96
Prob > chi2	0.00000	0.00000	0.00000	0.00000

For artists, the story for the regressions performed with only non-zero sales is similar. The results are listed in Table 18 and the coefficients on sales rank are graphed in Figure 25. All coefficients are statistically different from zero, $p<0.01$ and all R^2s are high. For all unit sales and non-digital sales, β_1 increases, or decreases in absolute value, each year for artists. This implies a long tail effect, but the year-to-year differences here are far less pronounced than above. This can be seen in Figure 25. Likewise, for digital units and tracks, the superstars trend is much less obvious. For artists the β_1 on **ln(Sales Rank)** does not change much for digital sales. Unbundled track sales for artists show no obvious trend.

Table 22: Regression Results for Non-Zero Artists, 2004-2008.

Regression results for artists with only *non-zero sales*: all unit sales, bundled digital sales, unbundled track sales, and non-digital sales. One asterisk (*) indicates significance at 5% and two (**) significance at 1%.

Artists, All Unit Sales

Non-Zero Sales Only	2004	2005	2006	2007	2008
β_0	19.67917	19.22728	18.64886	18.24244	18.08056
Std. Err.	0.14922	0.12332	0.09318	0.07696	0.07239
β_1	-2.87862**	-2.78440**	-2.63498**	-2.55351**	-2.48198**
Std. Err.	0.02648	0.02140	0.01562	0.01266	0.01166
R^2	0.94460	0.95550	0.96710	0.97400	0.97330

Artists, Bundled Digital Sales

Non-Zero Sales Only	2004	2005	2006	2007	2008
β_0	12.34504	13.60712	14.28978	15.08215	15.42584
Std. Err.	0.20338	0.15477	0.10522	0.08575	0.08930
β_1	-2.34946**	-2.33243**	-2.26225**	-2.29619**	-2.25552**
Std. Err.	0.04852	0.03268	0.01992	0.01542	0.01549
R^2	0.93760	0.94880	0.96400	0.97190	0.96400

Artists, Non-Digital Sales

Non-Zero Sales Only	2004	2005	2006	2007	2008
β_0	19.69318	19.14236	18.65203	18.15277	17.73801
Std. Err.	0.15206	0.12513	0.09900	0.08087	0.06988
β_1	-2.88943**	-2.78719**	-2.68022**	-2.60904**	-2.51903**
Std. Err.	0.02707	0.02190	0.01690	0.01366	0.01162
R^2	0.94370	0.95570	0.96660	0.97530	0.97880

Artists, Unbundled Track Sales

Non-Zero Sales Only	2004	2005	2006	2007	2008
β_0	18.49341	19.13396	19.93960	20.50949	21.38744
Std. Err.	0.25916	0.16426	0.13104	0.10924	0.11020
β_1	-3.18883**	-2.97902**	-2.93691**	-2.92860**	-3.01669**
Std. Err.	0.05832	0.03240	0.02402	0.01922	0.01896
R^2	0.93640	0.95610	0.96290	0.96960	0.96830

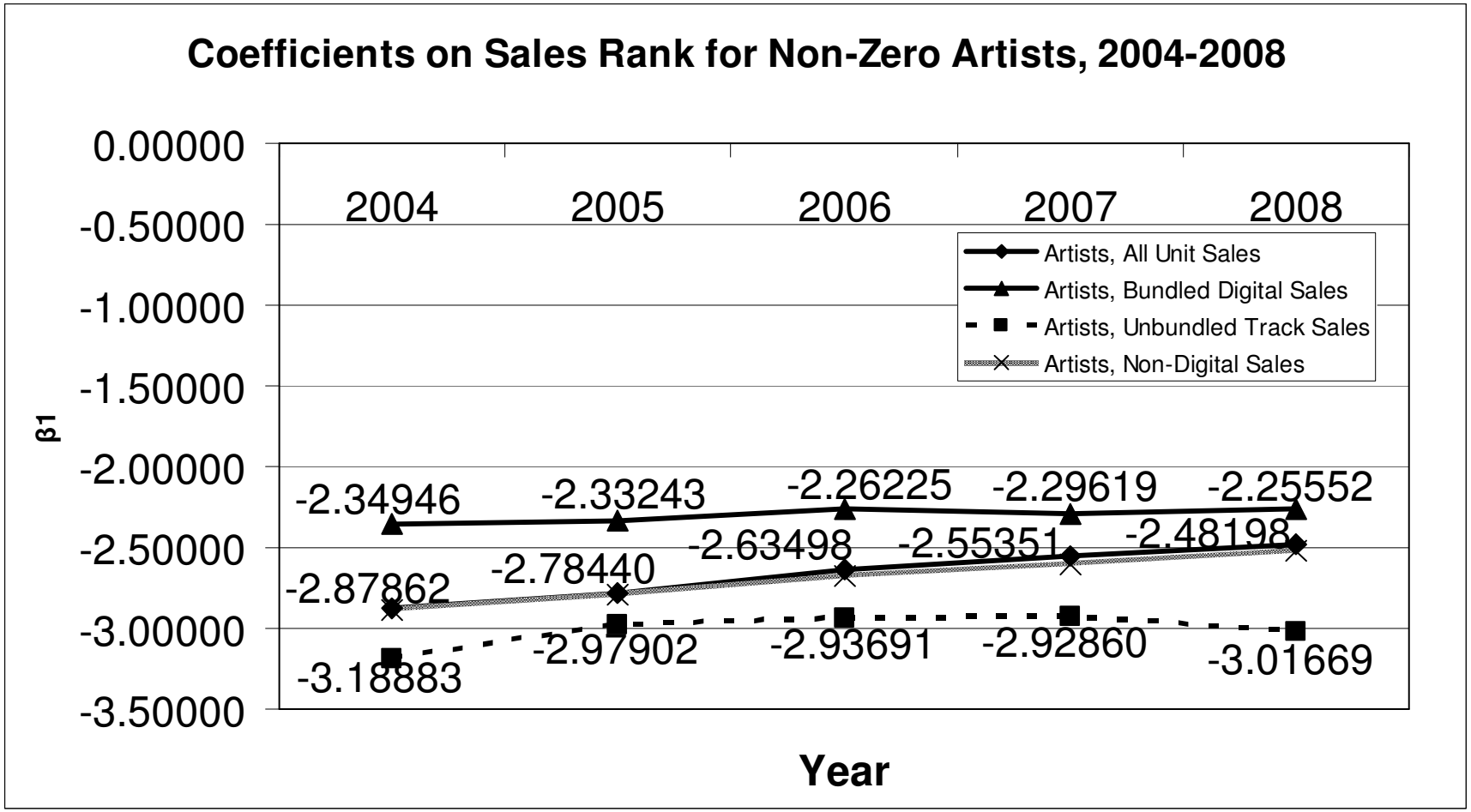

Figure 37: Coefficients on Sales Rank for Non-Zero Artists, 2004-2008.

Coefficients on sales rank for artists with only *non-zero sales* for all unit sales, bundled digital albums and singles, unbundled digital track sales, and non-digital unit sales for each year.

For artists with non-zero sales only, the tests for statistically significant differences between the β_1s between each pair of consecutive years are shown in Table 19. For all unit sales and non-digital sales, these are all highly significant, $p<0.01$. For bundled digital sales, the year-to-year differences are insignificant for 2004-2005, are significant at 5% for 2005-2006 and 2006-2007, and significant at 1% for 2007-2008. For unbundled track sales, the difference is significant at the 1% level from 2004-2005, insignificant in between and then significant at the 1% level again from 2007-2008. The same implications from the above regressions with sales of zero included still stand for all unit sales and non-digital sales when the zeros are removed from the regression: a significant long tail effect. For the two digital formats considered, the superstar effect is non-existent when only non-zero sales are considered in the regression.

Table 23: Test of Significance Between Years for Non-Zero Artists.

Results for the test of significance for the coefficient on sales rank (β_1) between each pair of consecutive years for artists with *non-zero sales*.

Artists, All Unit Sales

Non-Zero Sales Only	2005-2004	2006-2005	2007-2006	2008-2007
chi2(1)	27.22	56.59	41.53	109.66
Prob > chi2	0.00000	0.00000	0.00000	0.00000

Artists, Bundled Digital Sales

Non-Zero Sales Only	2005-2004	2006-2005	2007-2006	2008-2007
chi2(1)	0.08	4.43	4.94	18.1
Prob > chi2	0.77610	0.03540	0.02630	0.00000

Artists, Non-Digital Sales

Non-Zero Sales Only	2005-2004	2006-2005	2007-2006	2008-2007
chi2(1)	33.75	35.24	30.73	154.9
Prob > chi2	0.00000	0.00000	0.00000	0.00000

Artists, Unbundled Track Sales

Non-Zero Sales Only	2005-2004	2006-2005	2007-2006	2008-2007
chi2(1)	6.99	1.38	0.15	40.03
Prob > chi2	0.00820	0.23930	0.69890	0.00000

For titles, the story for the regressions performed with only non-zero sales is similar. The results are listed in Table 20. All coefficients are statistically different from zero, $p<0.01$ and all R^2s are high. For titles, the β_1 for all unit sales increases, or decreases in absolute value, each year. This implies a long tail effect, but the year-to-year differences are far smaller, as graphed in Figure 26. For digital units, the superstars trend is much dampened but still visible. For tracks, the superstars trend is present. Non-digital sales show no obvious trend. Again, the coefficients change much less from year to year, and trends are less prominent when only non-zero sales are considered.

Table 24: Regression Results for Non-Zero Titles, 2004-2008.

Regression results for titles with only *non-zero sales*: all unit sales, bundled digital sales, unbundled track sales, and non-digital sales. One asterisk (*) indicates significance at 5% and two (**) significance at 1%.

Titles, All Unit Sales

Non-Zero Sales Only	2004	2005	2006	2007	2008
β_0	21.47061	21.35977	21.48776	20.82042	20.12763
Std. Err.	0.11963	0.10765	0.09162	0.08235	0.07340
β_1	-2.58633**	-2.56329**	-2.55014**	-2.46444**	-2.34493**
Std. Err.	0.01739	0.01543	0.01286	0.01142	0.00998
R^2	0.90020	0.91070	0.92590	0.93130	0.93290

Titles, Bundled Digital Sales

Non-Zero Sales Only	2004	2005	2006	2007	2008
β_0	11.90182	13.57606	14.70792	15.38134	15.68703
Std. Err.	0.14887	0.11632	0.10281	0.08810	0.08300
β_1	-1.80366**	-1.90947**	-1.95696**	-1.97685**	-1.95991**
Std. Err.	0.02839	0.02012	0.01652	0.01359	0.01235
R^2	0.89690	0.91810	0.91760	0.92810	0.92360

Titles, Non-Digital Sales

Non-Zero Sales Only	2004	2005	2006	2007	2008
β_0	21.52278	21.40234	21.63546	21.00180	20.14196
Std. Err.	0.12286	0.11269	0.09992	0.09279	0.08144
β_1	-2.60200**	-2.58902**	-2.60989**	-2.55109**	-2.42254**
Std. Err.	0.01792	0.01630	0.01425	0.01317	0.01141
R^2	0.89800	0.90870	0.92290	0.92820	0.93400

Titles, Unbundled Track Sales

Non-Zero Sales Only	2004	2005	2006	2007	2008
β_0	19.66380	22.53804	24.03926	24.99087	25.89233
Std. Err.	0.32401	0.23665	0.18482	0.15333	0.14941
β_1	-2.54870**	-2.87007**	-2.98624**	-3.03299**	-3.09079**
Std. Err.	0.05857	0.03904	0.02887	0.02304	0.02190
R^2	0.75240	0.83480	0.87630	0.89870	0.89580

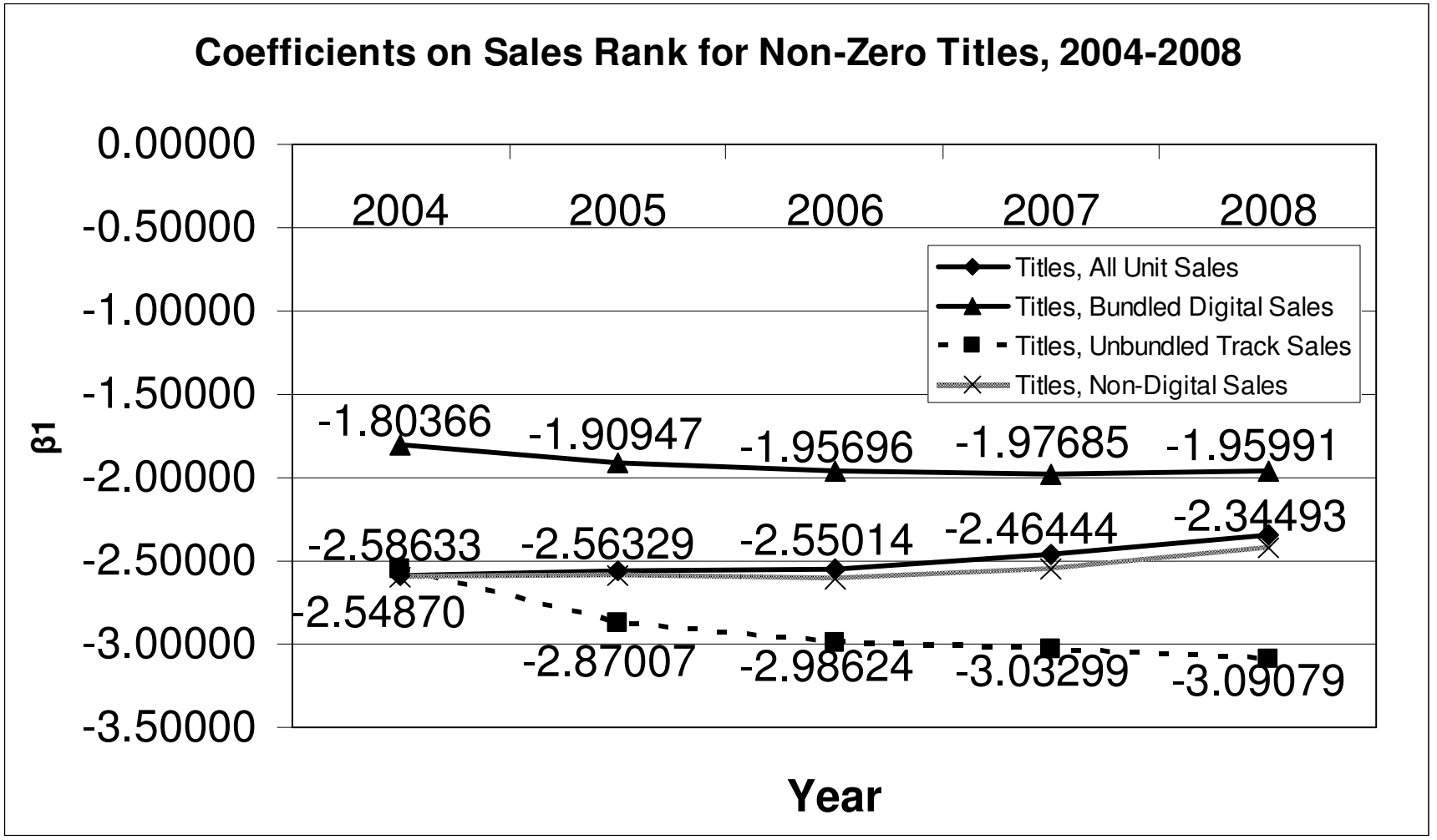

Figure 38: Coefficients on Sales Rank for Non-Zero Titles, 2004-2008.

Coefficients on sales rank for titles with only *non-zero sales* for all unit sales, bundled digital albums and singles, unbundled digital track sales, and non-digital unit sales for each year.

Considering titles with only non-zero sales, the test for statistically significant differences between the β_1s between each pair of consecutive years are shown in Table 21. For all unit sales, 2004-2005 is significant at 5%, but then insignificant from 2005-2006 until significant again at 1% from 2006-2007 and 2007-2008. For bundled digital sales, the year-to-year differences are significant at 1% for 2004-2005 and 2005-2006, then significant at the 5% level for 2006-2007 and 2007-2008. For non-digital sales, 2004-2005 is insignificant, but the following three pairs of years are significant at the 1% level. For unbundled track sales, all pairs of years are significant at the 1% level aside from 2006-2007. The same implications from the above regressions are relevant here for all unit sales and non-digital sales when the zeros are removed from the regression: a long tail effect, but much less consistent. For the two digital formats considered, the superstar effect is visible and statistically significant when only non-zero sales are considered in the regression.

Table 25: Test of Significance Between Years for Non- Zero Titles.

Results for the test of significance for the coefficient on sales rank (β_1) between each pair of consecutive years for titles with *non-zero sales*.

Titles, All Unit Sales

Non-Zero Sales Only	2005-2004	2006-2005	2007-2006	2008-2007
chi2(1)	6.4	2.29	347.69	323.63
Prob > chi2	0.01140	0.13010	0.00000	0.00000

Titles, Bundled Digital Sales

Non-Zero Sales Only	2005-2004	2006-2005	2007-2006	2008-2007
chi2(1)	11.4	9.53	3.94	4.34
Prob > chi2	0.00070	0.00200	0.04720	0.03710

Titles, Non-Digital Sales

Non-Zero Sales Only	2005-2004	2006-2005	2007-2006	2008-2007
chi2(1)	3.17	9.92	351.45	384.37
Prob > chi2	0.07520	0.00160	0.00000	0.00000

Titles, Unbundled Track Sales

Non-Zero Sales Only	2005-2004	2006-2005	2007-2006	2008-2007
chi2(1)	44.11	10.47	4.81	24.89
Prob > chi2	0.00000	0.00120	0.02830	0.00000

So, a superstar effect is evident in the digital formats: bundled digital albums and singles and unbundled tracks for both artists and titles. Consider this the 'digital superstar effect'. A long tail effect is evident in overall unit sales and non-digital sales for both artists and titles. Consider this the 'non-digital long tail effect'. However, when only non-zero sales are considered these trends are not as conspicuous, though they are still present. For non-zero sales only, there is asymmetry between artists and titles. For non-zero artist sales, the digital superstar effect is diluted while the non-digital long tail effect is still there. For non-zero title sales, the non-digital long tail effect is diluted while the digital superstar effect is still present. The conclusions gleaned from this chapter's data analysis are summarized in Table 23.

D. Summary of Results

For ease of comparison, Table 23 lists observations, the specific changes over time, and the implications of the specific changes for each hypothesis. Each hypothesis – the superstar, long tail, and hybrid, respectively – are restated. Each relevant observation and specific change pertaining to each respective hypothesis is grouped together under the hypothesis name and description in the order they are examined in this chapter.

Table 26: Synopsis of Data Analysis Conclusions.

Hypothesis 1.	**Superstar Hypothesis.**	
The superstar hypothesis is that sales have become relatively more skewed toward hits over time.		
Observation.	Specific Changes.	Implications for Hypotheses.
Percentage of sales accruing to each quantile of artists. (Table 4, Figure 11, Figure 12).	Market becomes more concentrated from 2004 to 2006.	Supports **superstar** hypothesis from 2004 to 2006.
Percentage of sales accruing to each quantile of titles. (Table 5, Figure 13, Figure 14).	Market becomes more concentrated from 2004 to 2006.	Supports **superstar** hypothesis from 2004 to 2006.
Percentage of sales accruing to each quantile of labels. (Table 6, Figure 15, Figure 16).	Market becomes more concentrated from 2004 to 2006.	Supports **superstar** hypothesis from 2004 to 2006.
Gini coefficients for artists. (Figure 17).	Market becomes more concentrated from 2004 to 2006. But, an earlier peak for digital, later peak for tracks.	Supports **superstar** hypothesis from 2004 to 2006.
Gini coefficients for titles. (Figure 18).	Market becomes more concentrated from 2004 to 2006. But, an earlier peak for digital, later peak for tracks.	Supports **superstar** hypothesis from 2004 to 2006.
Gini coefficients for labels. (Figure 19).	Market becomes more concentrated from 2004 to 2006. But, an earlier peak for digital.	Supports **superstar** hypothesis from 2004 to 2006.
Gini coefficients for non-zero artist sales. (Figure 20).	Market becomes more concentrated from 2004 to 2006. But, a later peak for tracks.	Supports **superstar** hypothesis from 2004 to 2006.
Gini coefficients for non-zero title sales. (Figure 21).	Market becomes more concentrated from 2004 to 2006. But, a later peak for tracks.	Supports **superstar** hypothesis from 2004 to 2006.
Gini coefficients for non-zero label sales. (Figure 22).	Market becomes more concentrated from 2004 to 2006. But, a later peak for non-digital.	Supports **superstar** hypothesis from 2004 to 2006.
Gini coefficients for non-	Market is more concentrated	Supports **superstar**

zero artists, titles, and labels. (Figure 20, Figure 21, Figure 22).	for all formats in 2008 than in 2004.	hypothesis from 2004 to 2008.
The coefficient on Sales Rank decreases each year for artists for digital units and for tracks. (Table 15, Figure 23).	Market becomes more concentrated each year.	Supports **superstar** hypothesis across sample period for **digital units** and for **tracks**.
The coefficient on Sales Rank decreases each year for titles for digital units and for tracks. (Table 17, Figure 24).	Market becomes more concentrated each year.	Supports **superstar** hypothesis across sample period for **digital units** and for **tracks**.
The coefficient on Sales Rank ***does not*** decrease each year for non-zero artist sales for digital units and tracks. (Table 19, Figure 25).	There is no perceptible change in market concentration by this measure for these formats.	***Does not*** support **superstar** hypothesis for **digital units** and for **tracks**.
The coefficient on Sales Rank decreases each year for non-zero title sales for digital units and for tracks. (Table 21, Figure 26).	Market becomes more concentrated each year.	Supports **superstar** hypothesis across sample period for **digital units** and for **tracks**.
Hypothesis 2.	**Long Tail Hypothesis.**	
The long tail hypothesis is that sales have become relatively less skewed toward hits and superstars over time.		
Observation.	Specific Changes.	Implications for Hypotheses.
Percentage of sales accruing to each quantile of artists. (Table 4, Figure 11, Figure 12).	Market becomes less concentrated from 2006 to 2008.	Supports **long tail** hypothesis from 2006 to 2008.
Percentage of sales accruing to each quantile of titles. (Table 5, Figure 13, Figure 14).	Market becomes less concentrated from 2006 to 2008.	Supports **long tail** hypothesis from 2006 to 2008.
Percentage of sales accruing to each quantile of labels. (Table 6, Figure 15, Figure 16).	Market becomes less concentrated from 2006 to 2008, but to a much smaller degree.	Supports **long tail** hypothesis from 2006 to 2008. But smaller in magnitude.
Gini coefficients for artists. (Figure 17).	Market becomes less concentrated from 2006 to 2008.	Supports **long tail** hypothesis from 2006 to 2008.
Gini coefficients for titles. (Figure 18).	Market becomes less concentrated from 2006 to 2008.	Supports **long tail** hypothesis from 2006 to 2008.
Gini coefficients for labels.	Market becomes less	Supports **long tail**

(Figure 19).	concentrated from 2006 to 2008.	hypothesis from 2006 to 2008.
Gini coefficients for non-zero artist sales. (Figure 20).	Market becomes less concentrated from 2006 to 2008.	Supports **long tail** hypothesis from 2006 to 2008. But smaller in magnitude.
Gini coefficients for non-zero title sales. (Figure 21).	Market becomes less concentrated from 2006 to 2008.	Supports **long tail** hypothesis from 2006 to 2008. But smaller in magnitude.
Gini coefficients for non-zero label sales. (Figure 22).	Market becomes less concentrated from 2006 to 2008.	Supports **long tail** hypothesis from 2006 to 2008. But smaller in magnitude.
Skewness declines each year for artists. (Table 9).	Market becomes less concentrated each year.	Supports **long tail** hypothesis across sample period.
The coefficient on Sales Rank increases each year for artists for all unit sales and non-digital sales. (Table 15, Figure 23).	Market becomes less concentrated each year.	Supports **long tail** hypothesis across sample period for **all unit sales** and **non-digital sales**.
The coefficient on Sales Rank increases each year for titles for all unit sales and non-digital sales. (Table 17, Figure 24).	Market becomes less concentrated each year.	Supports **long tail** hypothesis across sample period for **all unit sales** and **non-digital sales**.
The coefficient on Sales Rank increases each year for non-zero artist sales for all unit sales and non-digital sales. (Table 19, Figure 25).	Market becomes less concentrated each year.	Supports **long tail** hypothesis across sample period for **all unit sales** and **non-digital sales**.
The coefficient on Sales Rank increases each year for non-zero title sales for all unit sales and non-digital sales. (Table 21, Figure 26).	Market becomes less concentrated each year.	Supports **long tail** hypothesis across sample period for **all unit sales** and **non-digital sales.**
Hypothesis 3.	**Hybrid Hypothesis.**	
The hybrid hypothesis is a combination of the effects of the above two hypotheses. Also can be described as the 'shrinking middle class' hypothesis.		
Observation.	Specific Changes.	Implications for Hypotheses.
Inter-quartile measure decreases each year for	'Middle class' of artists is shrinking.	Supports **hybrid** hypothesis.

artists. (Table 9).		
Right tail inter-quartile measure decreases each year for artists. (Table 9).	'Middle class' of artists is shrinking.	Supports **hybrid** hypothesis.
Inter-quartile measure decreases each year for titles. (Table 10).	'Middle class' of titles is shrinking.	Supports **hybrid** hypothesis.

To reiterate, here are the hypotheses considered in light of the data analysis in this chapter:

Hypothesis 1. *There is a 'superstar' effect in the market for recorded music today. The superstar hypothesis* is that sales have become relatively more skewed toward hits over time. The market becomes more concentrated from 2004 to 2006 judging by the Gini coefficients and the percentage of sales accruing to each quantile of titles, artists, and labels. Regardless of the rise and slight fall from 2004 to 2008 in the Gini coefficients, when zero sales are dropped each Gini coefficient ends up higher in 2008 than it was in 2004. This increased concentration lends support to the superstar hypothesis. The regression results show the coefficient on Sales Rank decreases each year for artists and titles for digital units and for tracks. So, there is a superstar effect in digital formats. This is not the case, however, for non-zero sales aggregated by artist, where the regression results show no conspicuous trend.

Hypothesis 2. *There is a 'long-tail' effect in the market for recorded music today that helps new or insurgent artists increase sales at the expense of established or incumbent artists. The long tail hypothesis* is that sales have become relatively less skewed toward hits and superstars over time. The market becomes less concentrated from 2006 to 2008 judging by the Gini coefficients and the percentage of sales accruing to each quantile of titles, artists, and labels. The long tail effect is smaller in magnitude when only non-zero sales are analyzed. Skewness declines each year when sales are aggregated over both artists and titles, which supports the long tail

hypothesis. The regression results show the coefficient on Sales Rank increases each year for artists and titles for all unit sales and non-digital sales. So, there is a long tail effect in non-digital formats.

Hypothesis 3. *There is some combination of the 'superstar' and 'long-tail' effects. The hybrid hypothesis* is a combination of the effects of the above two hypotheses. The inter-quartile measure decreases each year for artists and titles and the right tail inter-quartile measure decreases each year for artists. These observations support the hybrid 'shrinking middle class' hypothesis.

V. Conclusions

The methodology used to answer the research questions does not rely on any interpretation of the superstar hypothesis which requires defining an objective measure of talent and/or making assumptions about maximizing consumption behavior. The premise in this research is that there has been an increase in supply facilitated by technological and institutional changes. Accordingly, the sales distribution profile is observed over time to see if sales concentration has increased or decreased. There is still lively debate about whether or not the long tail effect is a reality and whether or not firms should invest in the idea as a business model informing their inventory offerings. Few studies have been able to utilize Nielsen Soundscan sales data and still fewer have samples this large and with a window of time this recent.

A. Problem and Hypotheses

Established business models of incumbent producers have been circumvented by the Internet and the Digital Revolution. The recorded music industry is in free fall because of a market failure: recorded music is not a private good because rivalness and excludability have vanished. Property rights over recorded music – and other creative content in other industries – have been demolished. Many observers believe that nascent and less-known

artists and bands presently benefit from the new democratized and decentralized marketplace. This new economic reality is being touted because is fosters niches and little-known bands/artists and therefore meets more diverse consumer tastes because middlemen and intermediaries are out of the supply chain. But are consumers' tastes really that diverse? Should producers and retailers cater to the idea that the public demands more variety? After all, this requires changes in business strategies, which is costly and risky. Have these recent changes in technology altered the 'winner-take-all' society and provided opportunity for a 'new artistic middle class'?

Hypothesis 1. *There is a 'superstar' effect in the market for recorded music today. The superstar hypothesis* is that sales have become relatively more skewed toward hits over time. If the distribution of recorded music sales has shifted toward top quantiles over time at the expense of all the other quantiles, this hypothesis will be supported.

Hypothesis 2. *There is a 'long-tail' effect in the market for recorded music today that helps new or insurgent artists increase sales at the expense of established or incumbent artists. The long tail hypothesis* is that sales have become relatively less skewed toward hits and superstars over time. If the distribution of recorded music sales has shifted toward bottom quantiles over time at the expense of all other quantiles, this hypothesis will be supported.

Hypothesis 3. *There is some combination of the 'superstar' and 'long-tail' effects. The hybrid hypothesis* is a combination of the effects of the above two hypotheses. If the distribution of recorded music sales has shifted such that both top and bottom quantiles increase at the expense of the middle quantiles, this hypothesis will be supported.

B. Findings and Implications

Simply looking at sales trends over time, familiar patterns unfold. The decline in CD sales found in the sample is consistent with broad industry trends shown in Figure 4 from the RIAA. Digital unit sales, meaning bundled albums and singles sold via an online retail platform, have been on the rise. This trend is much touted in the industry as a success story. Unfortunately, this category of sales is comparatively small, and is not compensating for the drop in CD sales. The exploding numbers for unbundled track sales are quite dramatic. This says that consumers are more and more willing to pick and choose individual tracks instead of an entire album, if given the choice. For record companies and retailers, this change in availability of individual tracks is referred to as a shift from a 'pure bundling' strategy to a 'mixed bundling' strategy where the album and all individual tracks are available separately. Elberse (2009) finds that this strategy lowers revenues, though revenues are not observable in the data used here. The album concept may be disappearing, but another observation is that LP album sales are rising. This is a niche market for audiophiles who, among other attributes, prefer analog sound over digital sound and value the album art that comes with the record.

The relative shares of sales accruing to the different quantiles of artists, titles, and labels show a general trend of rising concentration – a superstar effect – for the first three years, and declining concentration – a long tail effect – in the final two years. The top 1% steals the lion's share of sales in each year for artists, titles, and labels. If artists, titles, and labels that sell zero are included, the Gini coefficients mirror this pattern of increasing and decreasing concentration. If only non-zero sales are included, the Gini coefficients for unbundled track sales and digital units, aggregated by artist and title, end up higher in 2008

than they were in 2004. This corroborates the digital superstar effect. The remainder of the formats, all unit sales and non-digital, for artists, titles, and labels show no obvious trend when only non-zero sales are included. In general, from 2006 to 2008, there is a slight decline in concentration visible in the Gini coefficients for artists, titles, and labels.

As far as the absolute levels of sales, the number of artists and titles who sell zero increases consistently each year. This probably reflects the increase in supply but also an increase in competition as the market is flooded with new artists/bands while demand for recorded music does not perceptibly shift or grow. The percentage of artists who sell one or more units remains about constant at 60%. The percentage of titles which sell one or more units declines slightly. This disparity between artists and titles is likely a result of mixed bundling because consumers are choosing individual tracks over album units over time. The album unit is aggregated as a title. For artists, skewness declines each year which implies less concentration. This is evidence for a long tail effect.

The regression coefficients provide evidence of a long tail effect in all sales and non-digital sales for artists and titles. On the other hand, there is evidence for a superstar effect in the digital formats: bundled digital units and unbundled digital tracks. So, there is evidence of a digital superstar effect and a non-digital long tail effect. This stands out because other researchers have found the opposite: a long tail effect in digital marketing channels when comparing the same products also sold in brick-and-mortar marketing channels (Brynjolfsson, Hu, and Simester 2007). It is likely that there are characteristics unique to recorded music when it comes to the consumption of cultural content akin to the snowball effect, which may accelerate with the new technological possibilities. For the regressions which include zero sales, all coefficients are significant and so are the tests for year-to-year differences. For the regressions which exclude sales of zero, all coefficients are still

significant, but many of the year-to-year differences are not statistically significant. The superstar and long tail effects are less obvious though still present when zero sales are chopped off. There is a differential effect visible in artists and in titles for non-zero sales only. For non-zero artist sales, the digital superstar effect is weakened while the non-digital long tail effect is still apparent. For non-zero title sales, the non-digital long tail effect is weakened while the digital superstar effect is still there. This is a matter of aggregation of sales over artists or over titles. It may be that many indie artists in the tail of the sales distribution have fans who prefer to purchase physical records while the fans of popular artists prefer digital content. This loyalty for artists does not show up in the individual titles. Perhaps popular artists in the head of the sales distribution have fans who purchase only certain superstar titles: albums and singles.

Illegal digital consumption has been expanding. Illegal digital consumption has been demonstrated to be a superstar market (Page and Garland 2009). Illegal downloads are dominated by the more popular artists/bands. Less popular artists are not hurt as much by illegal digital consumption as popular artists are (Blackburn 2004). Little-known and new artists often release their music online for free anyway for marketing purposes. Less popular artists also would likely benefit minimally from releasing their music on legitimate online retail distribution platforms. This is because their volume of sales is so much less that it is not worth the cost for the artist or the retailer. More popular artists, backed by record labels, may be more ready to embrace legitimate Internet distribution for this reason. So, record label backing could be affecting outcomes here. This may help to explain the digital superstars effect. However, there is evidence that the long tail effect is present in overall sales measures and in non-digital formats. This may be because indie artists may focus more on the album concept. Therefore, physical album sales may be preferred by these artists and

their fans and these remunerated sales show up in the long tail of the data set's distribution profile.

Most consumers may be illegally downloading due to income constraints and are maximizing their consumption of recorded music subject to this constraint. But other consumers feel the need to remunerate the artists/bands whose music they enjoy because of feelings of reciprocity. At the same time, many artists and their audiences, the world over, view making money as negative due to an ascetic aesthetic disposition (Bourdieu 1993, 1984). Indie rock, as a concept, revolves around the distinction of a lack of a mass audience or a lack of popularity. Popularity is considered 'vulgar'. In that way, its fans can have the benefit of an exclusive aesthetic awareness (Hibbet 2005). Economically, this is a voluntary constraint on upward mobility, which might otherwise be observed in the sales distribution. Interestingly, this anti-mass market attitude does provide some social/cultural excludability. This does not show up in the legitimate transaction data used for this study, but may be an interesting avenue for future research.

Recorded music has become non-rival because there is no scarcity. Recorded music has also become non-excludable because intellectual property rights are ever harder and costlier to enforce. A public good, which is the absence of rivalness and excludability, is a market failure. The long-tail-technological-institutional changes allow for non rivalness in inventory, which is economically positive for producers and consumers alike. The non rivalness in inter-individual consumption, i.e. copying and file sharing, is economically negative for popular content creators and owners because this decreases their revenues. On the other hand, it is economically positive for less-popular content creators and owners because they benefit from the free marketing and dissemination of their artwork. Digitization and file sharing also allows non-excludability to occur such that content creators

and copyright owners have lost much of the enforceability of their property rights. The end result is that recorded music prices are falling toward zero.

Consumption rivalness is highly unlikely to come back in the digital era, so excludability is the only strategy for copyright holders. Will Page, of PRS for Music, a copyright collecting society in the U.K., suggests that bundling is the likely best response. This is the case as long as the reserve price or willingness to pay for the bundled good or content access, i.e. an album or a music subscription service, is less dispersed than the reserve price or willingness to pay for the unbundled parts, i.e. individual tracks. The marginal cost of digital music has become insignificant, so the task for the suppliers is to maximize revenue given that consumer reserve prices differ for different tracks or individual components. This is illustrated in Table 24. What is shown is a case where the retailer is deciding between a strategy of strict bundling and a strategy of strict unbundling, and not mixed bundling. One consumer, Ted, values Track 1 more than Track 2. Another consumer, Laura, values Track 2 more than Track 1. First, individual tracks could be sold separately at $1.00 a piece. In this case revenue would be $4.00. Second, individual tracks could be sold separately at $1.20 a piece. In that case revenue would only be $2.20 because one of the two tracks is not worth the price to each respective customer. However, if both Track 1 and Track 2 were bundled together and sold at the combined price of $2.20, which corresponds to the combined willingness to pay for each customer, both would buy the bundle and revenue would be maximized at $4.40. In this way the willingness to pay for the bundle is less dispersed than the willingness to pay for the individual tracks.

Table 27: The Willingness to Pay for a Bundle is Less Dispersed than the Willingness to Pay for Unbundled Components.

The prices listed represent reserve prices for each person for the respective tracks. If all tracks were priced at $1.20 and sold separately, only two would sell: Ted would buy Track 1 and Laura would buy Track 2, and revenue would be $2.40. If all tracks were priced at $1.00 and sold separately, all four tracks would sell, and revenue would be $4.00. If Track 1 and Track 2 were bundled together and sold for $2.20 then revenue would rise to $4.40. Source: Page 2006.

	Track 1	Track 2
Ted	$1.20	$1.00
Laura	$1.00	$1.20

Elberse (2009) finds empirical evidence that the mixed bundling strategy lowers revenue compared with pure bundling. Mixed bundling, as mentioned earlier, is a case where the bundled album is available as a choice for the consumer but so are the individual unbundled tracks. Certain artists already prefer bundled albums for personal expression as an art form, but quite often retailers have bargaining power over artists and do not allow pure bundling strategies, e.g. iTunes. If the bundling of albums is less than successful, then bundling in the form of live performances and selling periodic access to bundled content through subscription services could reinstate some excludability. Artists and bands should have freedom in pricing albums and tracks as they see fit. An example of this where the artist is empowered to set price is CD Baby. CD Baby is both a record label and online distribution platform. This has potential to fight the general trend of falling prices of recorded music (Page 2006).

Given the tendency toward the public good status of recorded music, the notion of the 'tragedy of the anti-commons' is relevant (Heller 1998, cited in Page 2006). The tragedy of the commons (Hardin 1968, cited in Page 2006) is a situation where a scarce resource is over-utilized due to insufficient rights holders by individuals acting separately in their own

self-interest. The tragedy of the commons is caused by not enough excludability. The tragedy of the anti-commons, on the other hand, is a situation where many individuals acting separately in their own self-interest under-produce a scarce resource because there are too many rights holders. This is relevant to the recorded music industry in the context of the digital rights management (DRM) reaction to piracy as well as RIAA litigation. This enforcement response to the non-excludability problem of the public good is an attempt to make a resource scarce by enforcing a right of exclusion or property right in this case.

> "Because intellectual property is sharable […] and also because it's invisible, making it difficult to ascertain boundaries, there is a lot more difficulty in implementing a system of exclusive property rights than we have with physical property. And so there are dangers in trying to transport wholesale our system of physical property rights into intellectual property" (Posner 2004, 73).

To curtail the growth of a black market for illegal copies of digital music, the Recording Industry Association of America (RIAA) sued Napster in late 1999. The RIAA won the lawsuit, but this was only one weed out of the garden for the trade group. These developments beg the question: how far should a society go to protect intellectual property rights? One line of research that has drawn a great amount of attention is Boldrin and Levine (2004). These authors argue that society has gone too far to protect intellectual property rights because any such rights granted by the government only enhance monopoly power. Monopoly power, so the economic argument goes, artificially inflates prices, curtails supply of the good/service, and hampers further innovative activity. Further, they assert that patents and copyrights are unnecessary because a perfectly competitive market provides a sufficient incentive structure for fostering innovation. However, incentives do matter and they are going away. If a person desires to be an artist, she faces the very real risk that the time and effort invested in human capital, networking, and devotion to the art – which is

required to score gigs, perfect one's craft, and hone one's skills – could be subject to prohibitive opportunity costs.

Record labels are not investing in artists as much as they used to, and this is a reaction to the public good reality. But there is no scarcity, as noted before, because of the attribute of non-rivalness. However, as more people use a resource, the resource's value or usefulness may increase due to the 'comedy of the commons' (Rose 1986 cited in Page 2006). This may have positive spillover effects onto other revenue streams such as merchandise or concert sales.

Though the recorded music industry is becoming democratized and decentralized and artists are taking their music directly to consumers, it does not necessarily follow that intermediaries such as record labels and copyright collectives are going to become irrelevant. This is especially true if those artists want to make a living. Given the sales distribution profiles observed in this data, collecting societies, digital aggregators, and record labels, and others will still play an important role because these intermediaries are able to apply unique economies of scale (Page 2008). The extensive and flat tailed sales distribution that is observed above leads to a huge number of micro payments and granular transactions. These transactions originate from a wide range of diverse online platforms, such as YouTube, iTunes, and Rhapsody, and they accrue to a diverse pool of niche artists. This justifies the existence of collecting societies and copyright collectives which collect and distribute royalty payments to their members. Copyright collectives can find solutions to coordination problems, bring more artists into cooperation, and decrease transactions costs between rights holder and end users (Heller 1998, cited in Page 2006).

C. Limitations of the Study and Future Research

There are several limitations to this research which are important to discuss. There is no information on the type of retailer from which individual unit sales go through. So analysis of marketing channels is not possible. There is also no information on the attributes of consumers making the purchases. Though it can be reasonably assumed that there has been a rightward shift in supply, changes in supply and demand cannot be separated with this data set. Therefore the changes in the distribution could be due to changes in music consumers' preferences or retailers changing their inventory offerings. Nielsen also refused to give information on the number of titles and artists entering the marketplace each year. Having this information in the future would help to explain to what degree supply changes translate into sales distribution changes.

Unbundled digital tracks are of dramatic importance in this data set. However, due to contractual time limitations we were unable to visit the part of the Nielsen system which reports 'digital tracks' as a separate format. That is, the unbundled tracks analyzed here are unbundled from the albums and singles captured in the sample. But digital tracks as a separate format have their own product code and are tallied separately. Analyzing this format in relation to albums, singles, and unbundled tracks would be an interesting avenue for future research. More information about the characteristics of artists would benefit the analysis, but these were not made available by Nielsen. For example, any proliferation of niches, as Chris Anderson argues should exist, would be visible in the proliferation of genres and sub-genres.

VI. List of References

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

Anderson, Chris. 2004. The Long Tail. Wired Magazine, Issue 12.10 (October 2004), http://www.wired.com/wired/archive/12.10/tail.html (accessed 9/15/2008).

Anderson, Chris. 2006. *The long tail: why the future of business is selling less of more*. New York: Hyperion.

Audretsch, David B. Joseph Schumpeter and John Kenneth Galbraith: 2 sides of the same coin? *Working Paper*. March, 2008.

Blackburn, David. On-line Piracy and Recorded Music Sales. *Working Paper*. December, 2004.

Boldrin, Michele and David K. Levine. (2004). The Case Against Intellectual Monopoly. *International Economic Review*. Vol. 45. pp. 327-350.

Bourdieu, Pierre. 1993. *The Field of Cultural Production*. Columbia University Press.

Bourdieu, Pierre. 1984. *Distinction: A Social Critique of the Judgment of Taste*. Trans Richard Nice. Cambridge, MA: Harvard Univ. Press.

Brynjolfsson, Erik, Yu Hu, and Duncan Simester. 2007. Goodbye Pareto Principle, Hello Long Tail: The Effect of Search Costs on the Concentration of Product Sales. *Working Paper*, MIT Center for eBusiness, Cambridge, Massachusetts.

Brynjolfsson, Erik, Yu Hu, and Michael Smith. 2003. Consumer Surplus in the Digital Economy: Estimating the Value of Increased Product Variety at Online Booksellers. *Management Science*. 49(11) 1580-1596.

Brynjolfsson, Erik, Yu Hu, and Michael Smith. 2006. From Niches to Riches: Anatomy of the Long Tail. *Sloan Management Review*. 47(4) 67-71.

Chevalier, Judith and Austan Goolsbee. 2003. Measuring Prices and Price Competition Online: Amazon.com and Barnesandnoble.com. *Quantitative Marketing and Economics*. 1: 203-222.

Chung, Kee and Raymond Cox. 1994. A Stochastic Model of Superstardom: An Application of the Yule Distribution. *The Review of Economics and Statistics*, 76 (4): 771-775.

Connolly, Marie and Alan Krueger. 2006. Rockonomics: The Economics of Popular Music. In *Handbook on the Economics of Arts and Culture*. Ed. Victor Ginsburgh and David Throsby. Elsevier.

Crain, Mark and Robert Tollison. 2002. Consumer Choice and the Popular Music Industry: A Test of the Superstar Theory. *Empirica,* 29: 1-9.

Demers, Joanna. 2006. *Steal this music: how IP law affects musical creativity*. GA: University of Georgia Press.

Diplo [Video]. 2010. Retrieved 5/25/2010, from http://www.thecreatorsproject.com/creators/diplo

Elberse, Anita. 2009. Bye Bye Bundles: The Impact of the Unbundling of Music in Digital Channels. *Journal of Marketing* (forthcoming).

Elberse, Anita and Felix Oberholzer-Gee. 2008. Superstars and Underdogs: An Examination of the Long-Tail Phenomenon in Video Sales. *Working Paper*. July 1, 2008.

Fisher, Irving. 1925. *Mathematical Investigations in the Theory of Value and Prices*. New Haven: Yale University Press.

Frank, Robert H. and Philip J. Cook. 1995. *The Winner-Take-All Society*. New York: The Free Press.

Giles, David. 2006. Superstardom in the US popular music industry revisited. *Economic Letters*, 92: 68-74.

Gomes, Lee. 2006. Many companies still cling to big hits to drive earnings. *Wall Street Journal*, August 2, Eastern Edition.

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

Hamlen, William. 1994. Variety and Superstardom in Popular Music. *Economic Inquiry*, 32 (3): 395-406.

Hardin, Garrett. 1968. Tragedy of the Commons. *Science*. 162, 1243-1248.

Heller, Michael. 1998. The Tragedy of the Anticommons: Property in the Transition from Marx to Markets. *Harvard Law Review*. 111 (3): 621-688.

Hibbett, Ryan. 2005. What is Indie Rock? *Popular Music and Society*. 28:1, pp. 55-77.

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

Leeds, Jeff. 2005. The net is a boon for indie labels. *New York Times*. Dec. 27, 2005. Arts Section. (Accessed Online, 4/6/2009).

Leeds, Jeff. 2007. The New Deal: Band as Brand. *New York Times*. Nov. 11, 2007. Music Section. (Accessed Online, 5/31/2009).

MacDonald, Glenn. 1988. The Economics of Rising Stars. *American Economic Review*, 78 (Mar.): 155-167.

Marshall, Alfred. 1947. *Principles of Economics, 8th Ed.* New York: MacMillan.

Orlowski, Andrew. 2008. Chopping the Long Tail Down to Size, Wired's Advice Could Seriously Damage Your Business. *The Register*. Nov. 7, 2008. Music and Media Section. (Accessed Online, 5/25/2010).

Page, Will and Eric Garland. 2009. The Long Tail of P2P. *PRS for Music*. Issue 14. May 14, 2009. 1-8.

Page, Will. 2008. Can 1,000 Fans Replace the Music Business? *The Register*. March 28, 2008.

Page, Will. 2007. Is Live the Future of Music? *The Report*. Issue 181. November 29, 2007.

Page, Will. 2006. Is the Price of Recorded Music Heading Towards Zero? Paper presented at Transmission Conference, November 30 – December 2, in Edmonton, AB, Canada.

Phoenix [Video]. 2010. Retrieved 5/28/2010, from http://www.thecreatorsproject.com/creators/phoenix

Posner, Richard A. (2004). "The Importance of Property Rights in the Common Law Tradition". Property Rights: The Essential Ingredient for Liberty & Progress. *Economic Education Bulletin: Proceedings of a Conference Sponsored by the American Institute for Economic Research*. pp. 67-76. A.I.E.R., Great Barrington, MA: Nov. 4-5.

Rose, Carol. 1986. The Comedy of the Commons: Custom, Commerce, and Inherently Public Property. *The University of Chicago Review*. 53: 711-781.

Rosen, Sherwin. 1981. The Economics of Superstars. *The American Economic Review*. 71(5) 845-858.

Schumpeter, Joseph A. 1911. *Theorie der wirtschaftlichen Entwicklung*. Berlin: Duncker und Humblot. Cited in Audretsch, 2008.

Schumpeter, Joseph A. 1942. *Capitalism, Socialism, and Democracy*. New York: Harper and Row Publishers. Cited in Audretsch, 2008.

Spierdijk, Laura, and Mark Voorneveld. 2007. Superstars without talent? The Yule distribution controversy. SSE/EFI Working Paper Series in Economics and Finance, No. 658. December, 5.

Tschmuck, Peter. 2006. *Creativity and Innovation in the Music Industry*. The Netherlands: Springer.