The Diminishing Returns of Music in the Digital Era*

Liban Timir

April 18, 2024

This paper explores the dynamic between digital transformation and revenue models in the music industry, inspecting data from industry sales to dissect the economic consequences of the digital shift. It builds upon and broadens existing research, pinpointing a significant downturn in revenue per unit despite escalated consumption rates, with a focus on the pivotal years (early 2000s) marking the industry's digitalization. The paper reveals a nuanced relationship: the paradoxical increase in access to music paired with decreased monetary valuation. These findings underscore the profound influence of digital market forces on the music industry and highlight the critical need for adaptive economic strategies within the rapidly evolving digital landscape.

1 Introduction

In the digital dawn of the 21st century, the music industry has experienced an unprecedented transformation, ushered in by the proliferation of the internet and a resultant seismic shift in consumer behavior. Specifically, the introduction of file sharing services like Napster redefined the way music is distributed, valued, and consumed by effectively making music free to share, download, and listen to online. The Recording Industry Association of America (RIAA) sought to shut Napster down and it did within 2 years. RIAA may have won the battle against Napster but lost the war against the internet. Historically, the industry's financial success was predicated on physical sales, but the advent of the internet has catalyzed a pivot to digital formats — a pivot not without its economic consequences. While the democratization of music through digital platforms has expanded reach, it has also ushered in a complex debate about the monetary value of music, with implications for artists, producers, and the industry at large.

*Code and data are available at: https://github.com/libant/music-industry-analysis

This paper seeks to explain the economic narrative of the music industry's journey through the digital age. By analyzing a comprehensive dataset on music sales, we unravel how the industry's revenue per unit has been affected since the decline of physical media. Amidst the backdrop of booming digital consumption, we find an inverse trajectory in revenue, raising questions about the long-term sustainability of current digital revenue models.

We employed quantitative methods to scrutinize patterns in music sales, juxtaposing the eras of physical dominance with the digital revolution. Our findings present that while digital consumption has indeed skyrocketed, the economic value per unit of music has suffered a significant decline. The implications of this research are multifaceted — highlighting the critical need for innovation in monetization strategies for music in an increasingly digital world.

The structure of the paper is as follows: We commence with a detailed examination of the dataset, charting the industry's sales and revenue trends (Section 2). Subsequent sections present our analysis of the data (Section 4), revealing pivotal insights into the economic shifts within the industry. We then pivot to a discussion (Section 5), contextualizing our findings within the broader framework of digital economics and their relevance to stakeholders. Finally, we conclude by reflecting on the implications of our research, considering the broader paradigm of value creation and capture in the digital era, and proposing pathways for future research. The analysis was conducted using the statistical programming language R (R Core Team (2023)), utilizing packages such as rstanarm (Goodrich et al. (2022)), and additional papers such as " ".

2 Data

2.1 Source

The main data source is a dataset from the Recording Industry Association of America (RIAA) (Larxel and Bass (2019)), tracking recorded music revenues by format in the United States from 1973 to 2019. This dataset was generated by Matt Bass and has been adjusted for inflation, offering a standardized monetary comparison across five decades of music industry economics. Since the RIAA have yet to properly respond to requests for the use of their dataset, we made use of a Kaggle dataset from Kaggle user "Larxel" that cites the RIAA dataset.

Detailing yearly revenue across a spectrum of music formats, the RIAA dataset offers a comprehensive look at the industry's economic fluctuations, tracing the lineage from traditional mediums like LPs and cassettes to modern digital forms such as streams and downloads. This level of detail affords a nuanced understanding of changing consumer preferences and technological advancements that have influenced the music market in the United States.

The RIAA dataset was specifically chosen for its focus on the U.S. market and its longitudinal data, allowing for an in-depth analysis of market shifts over fifty years. Other global datasets

did not offer the same level of historical depth or the focused granularity required to assess the U.S. market's unique journey through the digital transformation. The RIAA dataset is composed of multiple variables that encapsulate different facets of the music industry. Variables such as "Format", "Year", and "Value" have been carefully vetted to ensure consistent and accurate representations of the industry's financial landscape over time. This thorough examination guarantees that the paper rests on a solid empirical foundation, capable of yielding substantial insight into the evolution of the music industry.

2.2 Measurement

2.2.1 Variables and Their Measurement

The key variables included in the Kaggle dataset and their respective measures are as follows:

1. Year: The year of the data entry, serving as a marker for trend analysis.
2. Format: The music distribution format, categorized into variaous phyiscal formats such as Vinyl, CD, and various digital formats like Streaming and Downloads.
3. Revenue: Measured in millions and adjusted for inflation to present a uniform value metric over time.

The inflation adjustment applied to the revenue figures is particularly noteworthy, as it allows for an equitable comparison of the industry's financial status year-over-year, accounting for the changing value of money. Bias should be taken into consideration as the method of revenue tracking has evolved with technology and industry standards.

2.2.2 Revenue by Format

Revenue, the central variable in the RIAA dataset cited in the Kaggle dataset, is measured in millions of U.S. dollars and has been adjusted for inflation to 2019 values. The RIAA has segmented revenue into various categories, each corresponding to a different format:

1. LP/EP
2. Vinyl Single
3. 8-Track
4. Cassette
5. Cassette Single
6. Other Tapes
7. CD
8. CD Single
9. SACD
10. DVD Audio
11. Music Video (Physical)

12. Download Album
13. Download Single
14. Ringtones & Ringbacks
15. Download Music Video
16. Other Digital
17. Kiosk
18. Paid Subscription
19. On-Demand Streaming (Ad-Supported)
20. Other Ad-Supported Streaming
21. SoundExchange Distributions
22. Synchronization

Each category represents a particular mode of music consumption, and the revenue figures associated with these formats are used to track shifts in consumer preferences and technological advancements. All physical formats and digital formats were cleaned and grouped together.

2.3 Data Characteristics

The RIAA dataset published by Larxel and Bass (2019) was cleaned to account for the irrelevence of the "number_of_records" column. Using the dataset, we were able to replicate all of the means and standard deviations of the music formats, as shown in Table 1. The summary statistics illustrate how the high standard deviations across the board reflect a wide variation in yearly revenue, which could be attributed to a certain format's lifecycle or other market changes over time. For example, 8-Track tapes, despite their irrelevance, have an intriguingly high mean revenue.

However, newer formats like digital downloads present lower mean revenues and relatively lower standard deviations, indicating a more consistent performance year-over-year but at a lower profitability scale. This could reflect market saturation (which is growing exponentially in the age of the internet), pricing strategies, or the competition from streaming services. The disparity between these formats highlights the evolution of music consumption, with digital formats overtaking physical mediums and reflecting changing consumer preferences and the influence of technological advancements such as social media and streaming services on the music industry's revenue streams.

Table 1: Summary Statistics of the Music Formats

format	mean_revenue	sd_revenue
8 - Track	511.177232	984.411165
CD	4733.432620	5645.536764
CD Single	45.346843	85.966719
Cassette	1403.643775	1919.217701
Cassette Single	126.793441	158.928097
DVD Audio	2.699241	3.632631
Download Album	523.063861	454.588843
Download Music Video	16.108696	14.523113
Download Single	956.128235	471.170965
Kiosk	2.843023	1.782577
LP/EP	905.310691	1966.291355
Limited Tier Paid Subscription	614.897193	228.885536
Music Video (Physical)	198.247984	213.153230
On-Demand Streaming (Ad-Supported)	452.028121	270.051745
Other Ad-Supported Streaming	214.019553	80.656749
Other Digital	19.121265	1.893653
Other Tapes	5.971933	17.806481
Paid Subscription	1417.139072	1827.921687
Paid Subscriptions	12.118509	17.023962
Ringtones & Ringbacks	299.985818	359.027230
SACD	3.602776	7.206144
SoundExchange Distributions	447.869189	359.379065
Synchronization	224.290481	32.118650
Vinyl Single	156.419683	268.091990

Figure 1 illustrates the year-over-year percentage change in revenue for various music formats, demonstrating the music industry's dynamic nature over time. For example, the sharp volatility observed in the cassette format's line is indicative of its rise and fall in popularity, reflective of technological obsolescence and subsequent market shifts. CDs exhibit less volatility but show a general decline in recent years, which correlates with the proliferation of digital music. The download single format and LP/EP show resurgence possibly due to niche market appeal and the revival of vinyl records among collectors and other enthusiasts. These revivals can be traced to social media trends as well, further propagating the internet's immense influence on the music industry. The paid subscription format presents a growing trend, emphasizing the industry's shift towards streaming services as the primary source of music consumption. This graph not only highlights broader trends in the industry, but it also lines up the death of physical media with the rise of the internet, which have reshaped the way music is distributed and monetized. The main concern of the contemporary music industry is also present in this

graph as the most consumed format of music (paid subscriptions) are declining in revenue year by year.

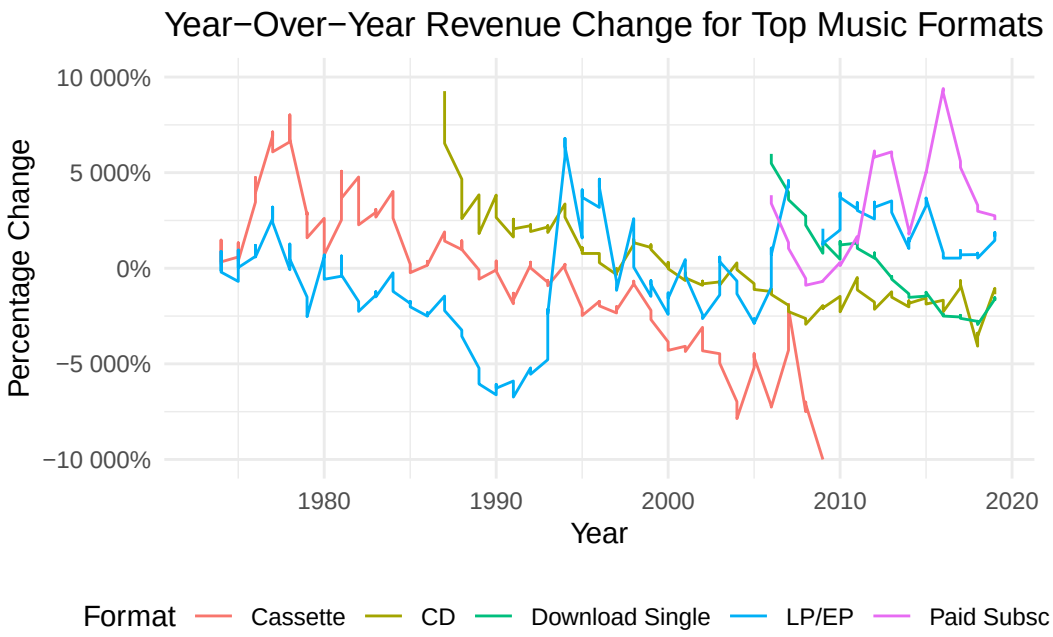


Figure 1: Revenue Change for the Top Music Formats

Table 2: Music Industry Average Revenue per Format

format	average_revenue
CD	4733.43
Cassette	1455.63
Paid Subscription	1417.14
8 - Track	1005.32
Download Single	956.13
LP/EP	905.31
Limited Tier Paid Subscription	614.90
Download Album	523.06
On-Demand Streaming (Ad-Supported)	452.03
SoundExchange Distributions	447.87
Ringtones & Ringbacks	299.99
Synchronization	224.29
Other Ad-Supported Streaming	214.02
Music Video (Physical)	198.25
Cassette Single	172.53
Vinyl Single	156.42
CD Single	46.83
Other Tapes	23.39
Other Digital	19.12
Download Music Video	16.11

Table 2 shows how in this sample, the most valuable format of music consumption in North America is the CD. This is because CDs are cheap to produce and there is a larger return on investment because they are physical format which have a larger cost than digital formats to the consumer. The least valuable format of music consumption in North America is the music video. This may be in part due to the fact that music videos are immensely costly and they have been slowly losing traction due to the rise in social media and streaming services.

Some of our data is of penguins (Figure 2), from Horst, Hill, and Gorman (2020).

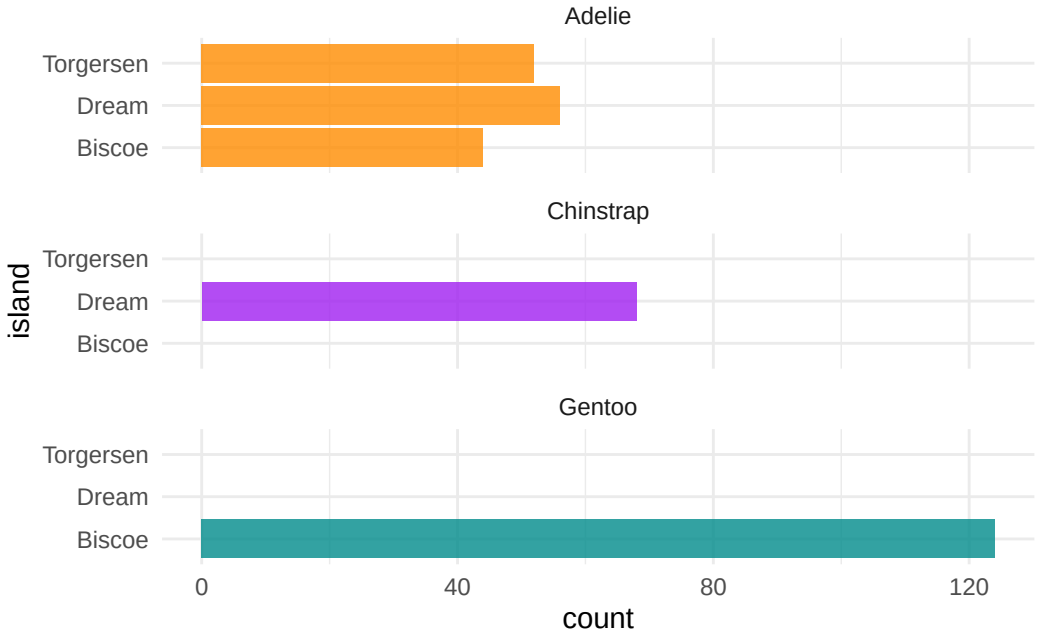


Figure 2: Bills of penguins

Talk more about it.

And also planes (Figure 3). (You can change the height and width, but don't worry about doing that until you have finished every other aspect of the paper - Quarto will try to make it look nice and the defaults usually work well once you have enough text.)

3 Model

The goal of our modelling strategy is twofold. Firstly, we seek to evaluate how different music formats contribute to overall industry revenue over time. This involves analyzing the changes in revenue streams from physical sales such as CDs and vinyl to digital formats like streaming and downloads. Secondly, we seek to use these insights to forecast future revenue trends for these formats.

In the section below, we outline the Bayesian analysis model utilized to examine the linear regression model of revenue generation per music format.

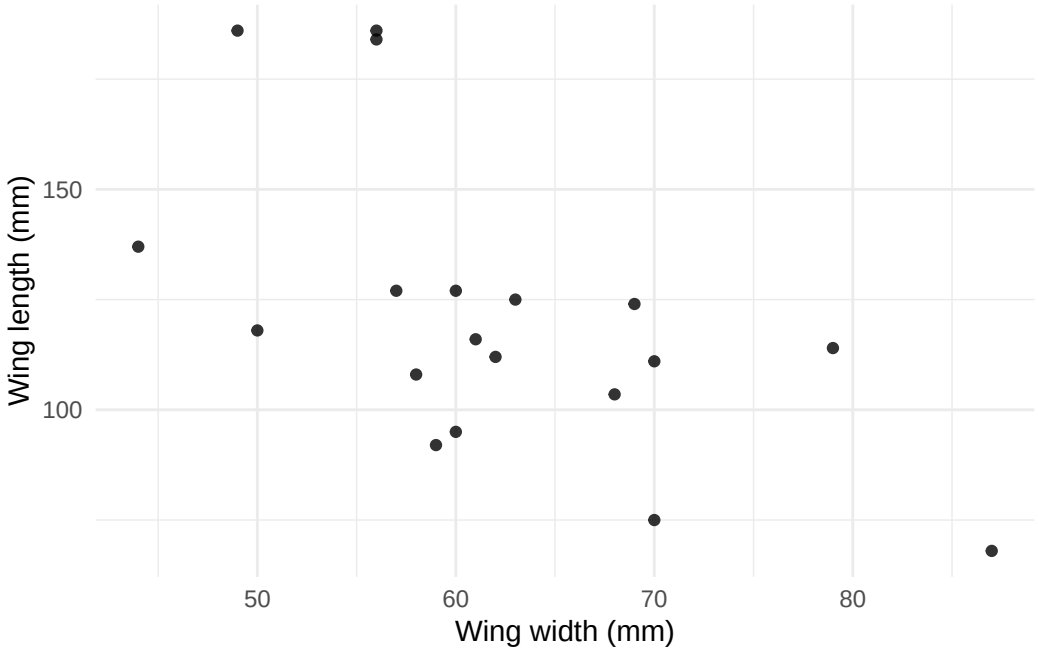


Figure 3: Relationship between wing length and width

3.1 Model set-up

Define y_i as the annual revenue from a given music format. Then β_i is the effect of physical sales, γ_i is the effect of digital downloads, θ_i is the effect of streaming subscriptions, and ϕ_i is the effect of ad-supported streaming. Let x_i be a vector containing the presence of each format in a given year.This is all measured in USD.

$$y_i|\mu_i, \sigma \sim \text{Normal}(\mu_i, \sigma) \tag{1}$$
$$\mu_i = \alpha + \beta_i x_{i1} + \gamma_i x_{i2} + \theta_i x_{i3} \tag{2}$$
$$\alpha \sim \text{Normal}(0, 2.5) \tag{3}$$
$$\beta \sim \text{Normal}(0, 2.5) \tag{4}$$
$$\gamma \sim \text{Normal}(0, 2.5) \tag{5}$$
$$\theta \sim \text{Normal}(0, 2.5) \tag{6}$$
$$\phi \sim \text{Normal}(0, 2.5) \tag{7}$$
$$\sigma \sim \text{Exponential}(1) \tag{8}$$

We run the model in R (R Core Team 2023) using the `rstanarm` package of Goodrich et al. (2022). We use the default priors from `rstanarm`.

3.1.1 Model justification

We expect a positive relationship between physical formats like CDs and vinyl records and revenue in the earlier years, reflecting the technology and consumer habits of those times. As we move into the digital age, we anticipate seeing a shift with digital downloads and streaming services gaining prominence and thus a positive relationship with revenue in the later years of the dataset.

Particularly, we predict that the rise of streaming services will correlate with a decrease in revenue from physical formats due to the convenience and accessibility of music online. Additionally, the advent of certain ad-supported streaming platforms might introduce a more complex relationship with revenue, potentially showing growth in consumer numbers but a slower rate of revenue increase due to its reliance on ad sales rather than direct consumer purchases.

Interactions between different formats are also of interest. For example, the coexistence of digital downloads and streaming services may reveal competitive or complementary dynamics. A competitive interaction might exhibit a negative relationship between the revenue of one format as the other rises, while a complementary interaction may not show a significant negative impact. These dynamics are crucial for understanding the evolution of music consumption and the financial viability of different formats in the music industry.

4 Results

Our results are summarized in Table 3.

5 Discussion

5.1 First discussion point

If my paper were 10 pages, then should be be at least 2.5 pages. The discussion is a chance to show off what you know and what you learnt from all this.

5.2 Second discussion point

5.3 Third discussion point

5.4 Weaknesses and next steps

Weaknesses and next steps should also be included.

Table 3: Explanatory models of flight time based on wing width and wing length

	First model
(Intercept)	1.12
	(1.70)
length	0.01
	(0.01)
width	−0.01
	(0.02)
Num.Obs.	19
R2	0.320
R2 Adj.	0.019
Log.Lik.	−18.128
ELPD	−21.6
ELPD s.e.	2.1
LOOIC	43.2
LOOIC s.e.	4.3
WAIC	42.7
RMSE	0.60

Appendix

A Additional data details

B Model details

B.1 Posterior predictive check

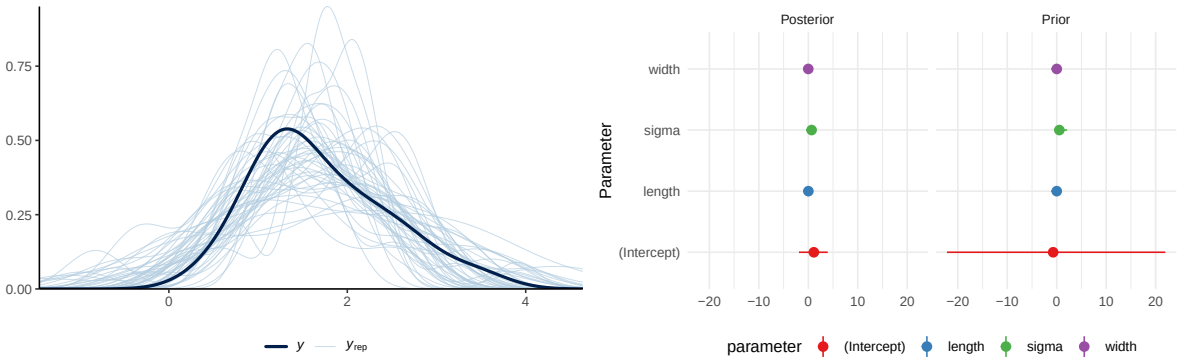
In Figure 4a we implement a posterior predictive check. This shows...

In Figure 4b we compare the posterior with the prior. This shows...

B.2 Diagnostics

Figure 5a is a trace plot. It shows... This suggests...

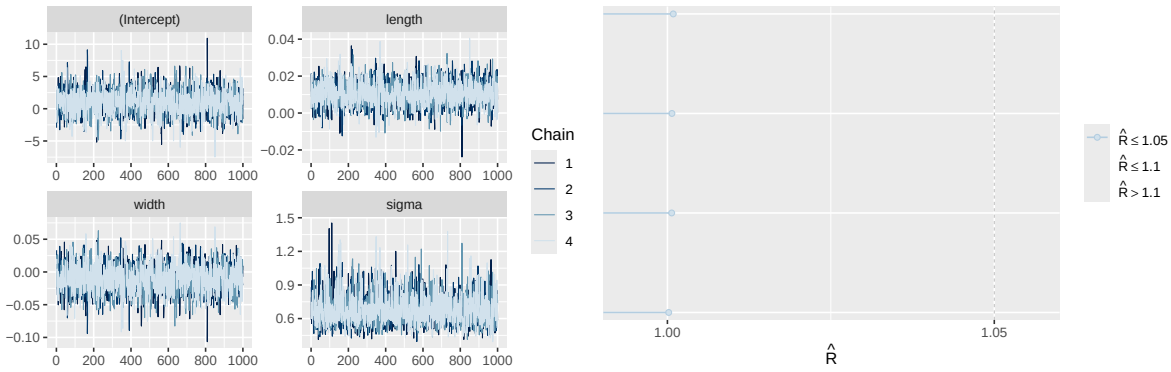
Figure 5b is a Rhat plot. It shows... This suggests...



(a) Posterior prediction check

(b) Comparing the posterior with the prior

Figure 4: Examining how the model fits, and is affected by, the data



(a) Trace plot

(b) Rhat plot

Figure 5: Checking the convergence of the MCMC algorithm

References

Goodrich, Ben, Jonah Gabry, Imad Ali, and Sam Brilleman. 2022. "Rstanarm: Bayesian Applied Regression Modeling via Stan." https://mc-stan.org/rstanarm/.

Horst, Allison Marie, Alison Presmanes Hill, and Kristen B Gorman. 2020. *Palmerpenguins: Palmer Archipelago (Antarctica) Penguin Data*. https://doi.org/10.5281/zenodo.3960218.

Larxel, and Matt Bass. 2019. "Music Industry Sales (1973 - 2019)." https://www.kaggle.com/datasets/andrewmvd/music-sales/data.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.