Playing to the Algorithm: How Spotify's Recommendations Shape Music Production*

Max E. Schnidman†

November 6, 2024; latest version available here

Abstract

I examine how recommender systems have influenced the music industry and shaped music production. Using a structural model of the recorded music industry, I analyze consumer behavior, platform recommendations, and rightsholder release decisions. I estimate a fixed cost of $170,000 for songs that enter Spotify's Top 200. Counterfactual analysis shows that with randomized recommendations, fewer songs would enter the market, reducing consumer welfare by 4%. The songs that do enter would be 33 seconds longer on average and more heterogeneously long. Popularity-based recommendations that do not account for individual taste would generate a superstar effect—increasing gross profit margins for songs that enter the market to 40%—but reducing consumer welfare by 13%. Although recommender systems have reduced overall variety in music, they have also enabled additional entry and increased consumer welfare.

Keywords: Recommender Systems, Economics of Platforms, Digital Economics, Music Economics

JEL Codes: D43, L15, Z11

*I extend my greatest thanks to my advising committee, Federico Ciliberto, Simon Anderson, and Julie Holland Mortimer. I also extend my thanks to Anton Korinek, Alex Mackay, Maxim Engers, Pete Troyan, Lidia Kosenkova, Gaurab Aryal, and seminar participants at the University of Virginia IO workshop. I acknowledge financial support from the Bankard Fund for Political Economy.

†University of Virginia, mes7jw@virginia.edu

1 Introduction

Recommender systems, which are designed to match consumers with products they will like, have transformed how consumers search for and acquire products. Such systems are prevalent in many online marketplaces, including Amazon, TikTok, and Netflix, and they have become a key feature of the digital music industry.[1] Music streaming platforms, on which consumers can access a vast catalog of music for a fixed monthly fee, have become the primary way consumers access music, with streaming accounting for 84% of the recorded music industry's $16bn revenue in 2023.[2] These platforms use recommender systems to generate playlists that surface music to users, and they are where users discover the majority of new music. I investigate how these recommender systems affect the music industry, and how they have shaped the sound of music since their introduction.

Recommender systems have generated significant regulatory and policy interest in recent years. Antitrust authorities have begun to investigate the effects of these systems on competition, and several pieces of legislation have been passed to regulate them. Examples include the Digital Markets Act and Digital Services Act in the EU and the US Department of Justice litigation against RealPage for algorithmic pricing collusion.[3] These systems also raise questions about artistic diversity and the long-term cultural impact of algorithmically driven music and cultural production. [4] This paper proposes a quantitative framework to analyze the effects of these systems on the music industry and estimate the welfare effects of these systems on consumers.

Recommender systems are a form of advertising for content on digital platforms, but unlike typical advertising, the producer does not actually purchase the advertisement.[5] These systems come with a number of economic trade-offs. Consumers can more easily find music they may like and discover new artists, and artists can reach a wider audience than ever before (Aridor and Gonçalves 2022). At the same time, platforms can use these algorithms to steer consumers toward profit-maximizing products, rather than products that consumers actually prefer (Reimers and Waldfogel 2023). These systems may also have inherent biases, resulting in recommendations that are not representative of the population or that are harmful to certain groups (Melchiorre et al. 2021).

I focus on the equilibrium effects of these systems—whereby producers respond to the

1. Amazon, TikTok, Netflix.
2. RIAA 2023 Year-End Music Industry Revenue Report.
3. Digital Markets Act, Digital Services Act, US Department of Justice, August 2024.
4. New Yorker, "Drowning in Slop."
5. Platforms do have sponsored recommendations, but Spotify, the platform I study, did not introduce these sponsored recommendations until after the timeframe of my data.

recommender system by changing their product design—and how these changes affect consumer welfare. To capture these equilibrium effects, I build a structural model of the recorded music industry to estimate the supply of and demand for recorded music on Spotify. This model has three sets of agents: consumers, Spotify, and rightsholders (producers). Consumers receive songs from Spotify's recommender system and choose whether to listen to them during their streaming session under a logit framework. Spotify's recommender system computes the probability that a consumer will listen to a particular song, based on the song's characteristics and the consumer's preferences, and delivers the song to the consumer. Rightsholders, such as record labels, choose whether to release songs to Spotify given the demand for the song, which is a function of the probability the recommender system surfaces the song and the consumer listens to it. They are forward-looking agents, looking to maximize expected profit, so they consider the future revenue the song generates when deciding whether to release it. In an oblivious equilibrium, rightsholders release songs so long as the expected revenue exceeds the fixed cost of release, whose distribution I estimate as a function of song characteristics and label fixed effects.

To estimate this model, I use three sources of data: the Music Streaming Session Dataset (MSSD), data scraped from Spotify Charts, and the Spotify API. The MSSD contains 160m consumer-level streaming sessions from July to September 2018. These streaming sessions include song characteristics, consumer characteristics, length of the listen (binned), and whether they got the song from a recommender system or other sources. Spotify Charts is a website reporting the daily top 200 songs on Spotify for every country in which they operate. It also includes stream counts and the Spotify ID for each song on the chart. The Spotify API allows me to query the song characteristics of each song on Spotify Charts.

I find that song characteristics, such as length, tempo, and danceability, have changed significantly since 2010. Using reduced-form analysis, I estimate that the introduction of streaming services and recommender systems correlate to a 40-second decrease in the average length of songs on Billboard's Hot 100. Music industry executives have also confirmed that they have changed the kind of music they release to better fit the recommender system's objectives (e.g., shorter, more danceable songs).

Using my structural model, I estimate a gap between consumer demand and recommender systems, driven by differences in their preferences, and find that producers respond to this gap by targeting the recommenders' objectives jointly with consumer preferences. For example, although consumers are likelier to listen to longer songs, the recommender system is likelier to surface shorter songs and producers respond by releasing shorter songs. I also estimate the fixed cost of releasing a song into the Top 200 on

Spotify at $170,000.

My counterfactual analysis focuses on changing Spotify's recommender system to see how it has affected song characteristics. I first impose random recommendations, as a proxy for no recommendations, or a naive consumer search process. I find that in the absence of recommender systems, songs are on average 33 seconds longer, more heterogeneous, and less profitable. As a result, fewer songs would be released, and consumer welfare is 4% lower than in the status quo. I also consider a recommender system based exclusively on song popularity, as a proxy for a ban on personalized recommendations, or a simulation of the digital music storefronts of the 2000s (e.g., iTunes). Such a recommender system would generate a superstar effect, which increases gross profit margins for songs that enter the market to 40%, but reduces consumer welfare by 13%. In this counterfactual, songs are 12 seconds shorter on average, and more danceable songs are likely to be released. This suggests that Spotify's recommender system has indeed changed the sound of music, and that while these changes have reduced the variety of music available to consumers, they have also increased both the quantity of songs and consumer welfare.

These results have significant economic and policy implications for the music industry and digital platforms more broadly. Economically, I find that recommender systems have reshaped the music production landscape by influencing not only consumer choices but also the creative decisions of artists and record labels. This shift has led to increased efficiency in matching consumer preferences with musical output, which potentially drives higher revenues and more targeted content creation. From a policy perspective, this research suggests that recommender systems may warrant regulatory scrutiny. Although the study indicates overall positive effects on consumer welfare, these systems can also drive concentration on digital platforms and in the music industry. Policymakers may need to balance the benefits of increased efficiency and consumer satisfaction against concerns about market power, artistic freedom, and cultural diversity. Furthermore, the study's methodology could inform future antitrust analyses and policy decisions regarding digital platforms and their recommendation algorithms across various industries.

The paper proceeds as follows. Subsection 1.1 places this paper in the context of the literature and identifies the contribution. Section 2 provides the background for the recorded music industry, describes the industry structure, including music characteristics, and provides reduced-form analysis of how technological changes have affected song characteristics in order to motivate the structural model. Section 3 describes the data and provides some descriptive analysis. Section 4 details the structural model of music streaming and describes the oblivious equilibrium in which rightsholders release music. Section 5 explains the estimation strategy. Section 6 provides and discusses the estimates of demand parameters, recommender system parameters, and fixed costs. Section 7 conducts several counterfactual analyses by modifying the recommender system to observe how equilibrium song releases change, and Section 8 concludes.

1.1 Literature Review

This paper contributes to several strands of the economics literature. First, it contributes to the literature on the economics of music by developing a structural model of the music streaming industry. Other works have analyzed the impact of Spotify on the industry. For instance, Aguiar, Waldfogel, and Waldfogel (2021) use reduced-form analysis to identify bias in the rankings of songs on Spotify's New Music Friday playlist. They find that higher-ranked songs tend to perform better after placement on the playlist, which suggests that curators are looking to maximize streams for their playlist. They also find that the curators of this playlist tend to favor songs by women and from independent labels, because they rank higher than their post-placement performance would suggest. Benner and Waldfogel (2016) use a difference-in-differences design to estimate how the digitization of recorded music has affected the release strategy of record labels. They find that, after digitization, major labels both release fewer albums and become more reliant on previously successful artists; conversely, independent labels release more albums. I extend these papers by applying these insights on Spotify playlists and digitization and embedding them in a structural model of the industry. This paper also builds on Aguiar and Waldfogel (2018), who develop a structural model of the digital music industry. They model consumer demand for digital music across countries and estimate the fixed cost of entry under three scenarios: perfect quality foresight, no quality foresight, and imperfect quality foresight, in which firms know their songs' quality, with some forecasting error. They estimate this fixed cost as the expected revenue of the worst-performing song, and find that the fixed cost is higher when rightsholders have no quality foresight. Their counterfactual analysis finds that tripling the number of songs available to consumers under imperfect foresight adds nearly 20 times as much consumer surplus as doing so under perfect foresight. I extend this model to the music streaming industry by modifying the choice structure to reflect the streaming industry, incorporating a recommender system in the model, and introducing forward-looking rightsholders.

Second, I contribute to a growing literature on recommender systems in economics. Bourreau and Gaudin (2022) use a Hotelling model of music listening with a recommender system and a digital platform that hosts both songs. They find that the platform uses the recommender system to drive consumers to songs with lower royalty rates, even if they are further from the consumer's ideal song. Aridor and Gonçalves (2022) similarly embed recommender systems in a theoretical model of digital platforms. They focus on the effect of these systems when the platform competes with its sellers (i.e., acts as a hybrid). They find that the platform uses the recommender system to steer consumers toward its own products, and that this can reduce consumer welfare through the foreclosure of third-party sellers. They also find that policy remedies are ambiguous in their effects, and that they can reduce consumer welfare if they are not carefully designed. I extend these analyses to an empirical model of the music industry and focus on how these systems affect producer product decisions. Melchiorre et al. (2021) introduce a large-scale dataset of music listening from Last.FM, a scrobbling service, and they use these data to investigate how several algorithms may exhibit gender bias. They find that significant disparities exist in recommendations with respect to certain gender groups. Aridor et al. (2023) conduct a field experiment to determine whether recommender systems affect consumption, using the recommendation service MovieLens. They find that recommender systems increase consumption beyond just the exposure provided by the recommendation. They also induce consumers to acquire additional information beyond what the recommendation provides. I apply their experiments to a structural model of the music industry.

Finally, I contribute to the literature on digital platforms and intermediation. Recent work in this area has focused on the role of platform exclusives and the possibility that these platforms can bias search and recommendation results toward certain profit-maximizing products, at the expense of consumer welfare. Lee (2013) constructs an empirical model of the video game industry that focuses on the role of exclusive games on console platforms. He finds that in the absence of exclusivity agreements, both console sales and consumer welfare would be higher, but only the incumbent console manufacturer would benefit from the absence of such agreements. I extend his model of game production to the music industry, and build on his use of first-order Markov processes to model firm dynamics. Reimers and Waldfogel (2023) develop an equilibrium framework to develop a workable definition of platform bias. Their model posits a welfare frontier for platforms, which is a weighted sum of consumer surplus and platform profits. They then test for biased rankings (recommendations) on the platform by evaluating whether the platform is on the frontier. They illustrate the approach by estimating the amount of bias in a structural model of Amazon and Expedia and find that both platforms are off the frontier. Aguiar and Waldfogel (2021) estimate the effect of including a song on a Spotify playlist using a regression discontinuity and instrumental variable design. They find that being included on a playlist significantly increases a song's eventual streams. I build on this work by incorporating algorithmic playlists in my model of the music industry.

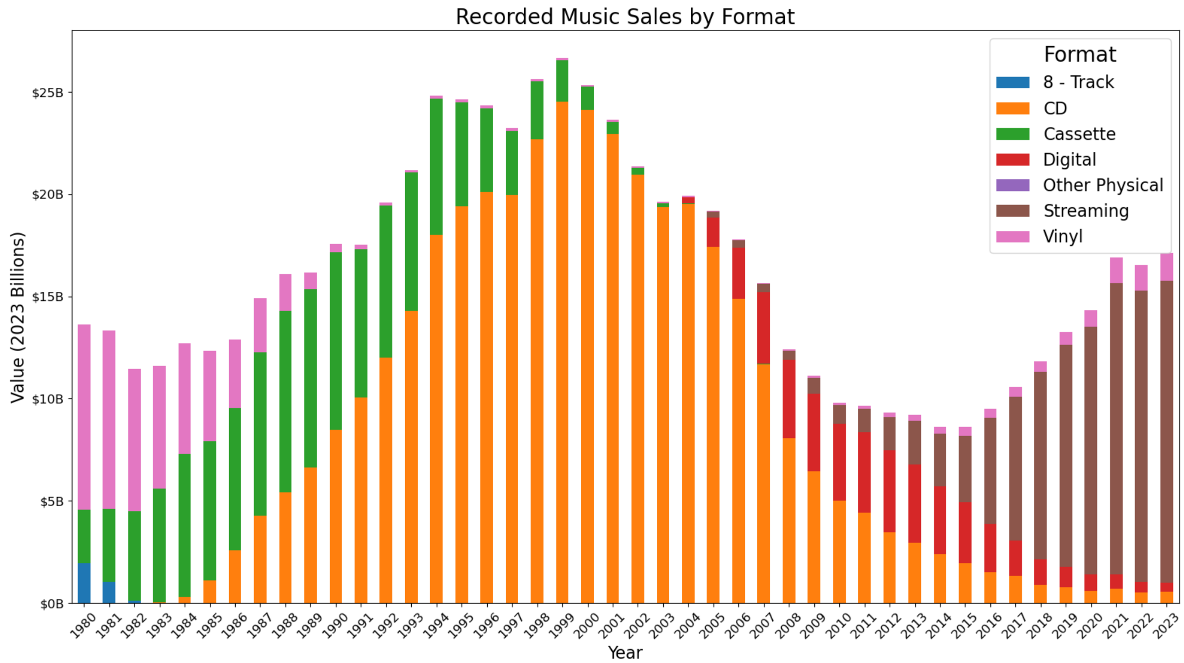

Figure 1: Real Revenue of the Recorded Music Industry, 1990-2023

Music revenue increased in real terms throughout the 2010s, with streaming services representing a growing share of revenue. Source: RIAA

2 Background and Industry Structure

2.1 Background

Technological changes have revolutionized the music industry over the last thirty years, as evinced by their fall and rise in real revenue in figure 1.

Growing access to the internet in the 1990s made it easier for consumers to digitally copy and share music, which led to the creation of illicit file-sharing services, such as Napster and Limewire, in the late 1990s. As these services spread in the early 2000s, revenues for the recorded music industry declined, which industry executives attributed to these services. Legal challenges brought by the Recording Industry Association of America resulted in the closure these services in the 2000s, but many copycat services emerged in their place.

To take advantage of the market for digital music and to support its iPod music players, Apple launched the iTunes store in 2003. iTunes made it easy for consumers to legally purchase digital music at low prices (99 cents per song). To address concerns about piracy, Apple made it difficult to share music sold on its platform, and designed its files such that they could only be played through iTunes or listened to on iPods.[6] Additionally, Apple negotiated a revenue-sharing deal with labels, giving them 30% of the revenue of every sale on iTunes, setting a precedent for revenue-sharing arrangements on digital platforms for the next two decades. While other digital companies attempted to launch their own music platforms (e.g., Google, Microsoft), none of them reached the level of financial success or cultural impact as iTunes. iTunes also broke up the album format, allowing consumers to purchase individual songs, rather than entire albums, another important precedent for streaming services.

In the late 2000s, some companies (e.g., Yahoo, Microsoft) began to experiment with streaming services, which provided consumers with on-demand access to an entire library of music for a subscription fee.[7] Such services did not see widespread acceptance until the early 2010s, when Spotify launched in the U.S. Spotify combined a 15,000,000-song library and an accessible two-tiered plan that included a free, ad-supported tier, and a paid, ad-free subscription tier. The subscription fee was waived for the first six months after launch.[8] Streaming made piracy much more difficult than copying digital downloads from iTunes or other digital platforms, because the service relied on streaming music from a centralized server.[9]

Spotify and similar streaming services (e.g., Apple Music, YouTube Music) proved incredibly popular, and helped to reverse the decline in the recorded music industry. Today, these services have become the primary way that consumers access music, with streaming accounting for 84% of the industry's revenue in 2023 (Figure 1).

2.1.1 Music and its Characteristics

Recorded music is the uniquely arranged combinations of sounds and vocals typically recorded in a studio. As a product, recorded music exists along numerous dimensions: length, chords, pitch, beats per minute, vocals, choices of instruments, etc. This results in infinitely many possible forms of music, ranging from the traditional (e.g., Beethoven's Ninth Symphony) to the esoteric (e.g. John Cage's 4'33"). Many of these dimensions are continuous, making it possible to use them as characteristics in a model of consumer preferences. (Lancaster 1966). In addition to the classical characteristics from music

6. https://www.engadget.com/2013-04-29-the-itunes-influence-part-one.html

7. https://www.thurrott.com/music-videos/groove-music/6033/microsoft-is-finally-retiring-zune-zune-music-pass

8. https://www.npr.org/sections/therecord/2011/07/14/137842612/spotify-has-arrived-stateside-heres-what-you-need-to-know, https://www.theverge.com/2012/1/6/2688250/spotify-free-account-restriction-10-hours-per-month

9. Amusingly, Spotify initially used pirated music before its agreements with record labels (Eriksson et al. 2019)

Name	Artist	Duration (min)	Tempo (BPM)	Key	Danceability	Energy	Speechiness	Valence
Sympathy for the Devil	Rolling Stones	6.3	116	A	0.7	0.67	0.21	0.56
Bohemian Rhapsody	Queen	5.9	71	C	0.41	0.40	0.05	0.22
Sweet Dreams	Eurythmics	3.6	125	C	0.69	0.71	0.03	0.88
Bad Romance	Lady Gaga	4.9	119	C	0.7	0.92	0.04	0.71
My Universe	BTS, Coldplay	3.8	105	A	0.59	0.7	0.04	0.44

Table 1: Examples of Song Characteristics

The first set of characteristics come from music theory (e.g., tempo, key), while the second set come from machine learning models (e.g., danceability, energy, valence). Source: Spotify API

theory (e.g, key, tempo, time signature), I include characteristics from machine learning models (e.g., danceability, energy, valence) in my model. I include descriptions of these characteristics in the Appendix (see Table 20).

Table 1 presents some examples of characteristics for popular songs.

Recently, cultural critics have observed a decrease in pop song length over the last twenty years, alongside a decrease in title length and an increase in lyric density.[10] In Figure 2, I plot the average length of songs on Billboard's Hot 100, by release year, finding that the average length of songs has been decreasing over time, with a noticeable acceleration in the 2010s.

To augment this, I conduct a reduced-form analysis of songs on Billboard's Hot 100 to confirm these trends. Specifically, I estimate the correlation between the introduction of new music formats and song duration. My regression equation is the following:

$$\begin{aligned} \text{Duration}_j &= \beta_0 + \beta_1 \mathbb{1}\{\text{Vinyl}\}_t + \beta_2 \mathbb{1}\{\text{Cassette}\}_t + \beta_3 \mathbb{1}\{\text{CD}\}_t \\ &= \beta_4 \mathbb{1}\{\text{Digital}\}_t + \beta_5 \mathbb{1}\{\text{Streaming}\}_t + \beta_6 \mathbb{1}\{\text{Recommenders}\}_t + \epsilon_j \end{aligned} \tag{1}$$

Each independent variable is an indicator variable for whether the particular format or technology was available at the time of the song's release. Table 2 reports the results of this regression.

These results are all statistically significant at the 1% level, and are negative for both the introduction of streaming services in the US in 2011 (as exemplified by Spotify), and the deployment of recommender systems on Spotify in 2015 (after their acquisition of Echo Nest). Combined, the introduction of these technologies are correlated with a 40-second decrease of average song length for songs that make it to Billboard's Hot 100, when comparing songs released in 2018 to songs released in 2010. This analysis is consistent with anecdotal evidence of changes in songs since the introduction of streaming services, but it

10. https://michaeltauberg.medium.com/music-and-our-attention-spans-are-getting-shorter-8be37b5c2d67

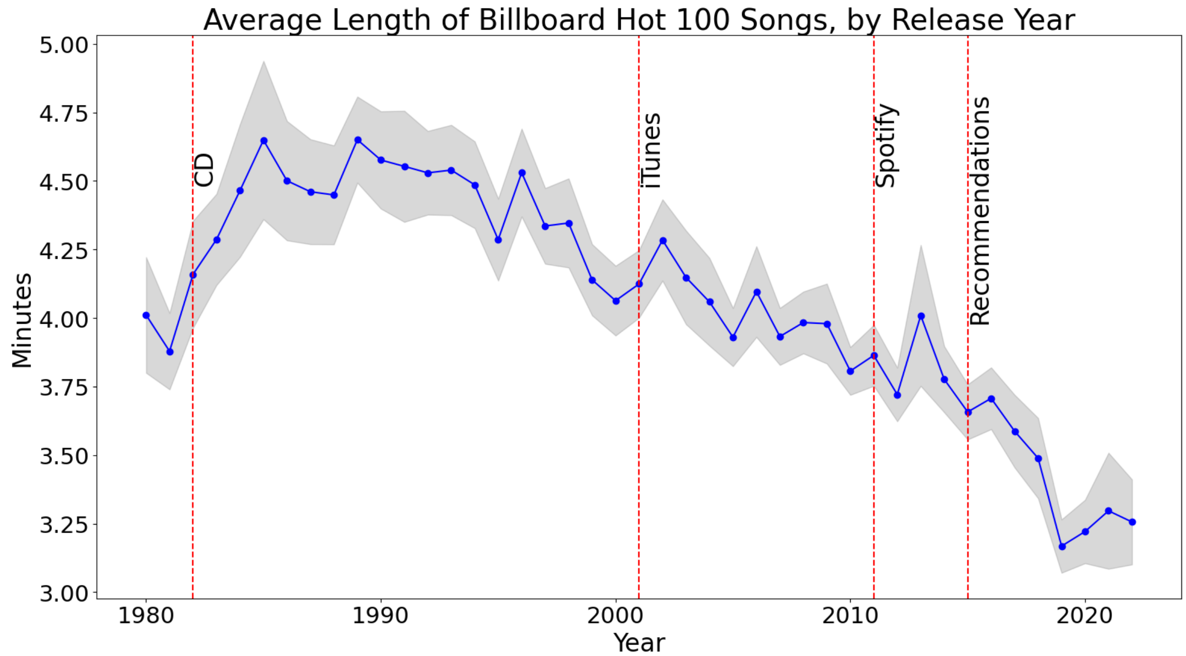

Figure 2: Average Song Duration on Billboard's Hot 100, 1990-2022

The average length of songs on Billboard's Hot 100 has been decreasing over time, with a noticeable acceleration in the 2010s. Source: Billboard

does not establish a causal relationship, or the mechanism by which these changes occur. For that I construct a structural model of the industry, whose agents and relationships I describe in the following subsection.[11]

2.2 Industry Structure

I group the recorded music industry into four sets of agents: artists, rightsholders, streaming platforms, and consumers. Figure 3 maps out the relationships between these agents.

Beginning on the left, artists create music, either by themselves or in contract with a record label, who serves as a rightsholder. An artist on contract with a rightsholder typically receives an advance and production assistance in exchange for ownership over the music they create. Artists also receive a share of the revenue (royalties) from the music they create, as negotiated with the rightsholders.[12] The market for artists is highly diffuse, with tens of thousands of artists working on music each day, competing not just with

11. In Appendix 8, I examine whether consumer preferences have changed over time, and whether these preferences are driving the changes in song length.

12. Song Royalties are an incredibly complex area of law, which I simplify for the purpose of this analysis by focusing on the payments between rightsholders and platforms. For a more detailed explanation, see https://www.royaltyexchange.com/blog/music-royalties-101-intro-to-royalties

	Dependent variable: Duration (m)
Recommendations	-0.394***
	(0.081)
Streaming Services	-0.207***
	(0.047)
Digital Sales	-0.343***
	(0.052)
CD	0.882***
	(0.141)
Cassette	0.824***
	(0.142)
Vinyl	-0.349***
	(0.049)
Intercept	3.023***
	(0.020)
Observations	6879
N. of songs	6276
N. of years	84
R^2	0.237
Residual Std. Error	0.522 (df=6872)
F Statistic	355.734*** (df=7; 6872)
Note:	*p<0.1; **p<0.05; ***p<0.01
Standard Errors clustered at the year level	

Table 2: Reduced Form Regression Results

The introduction of streaming services correlates to a 12-second decrease in song length, while the introduction of recommender systems correlates to a 24-second decrease in song length.

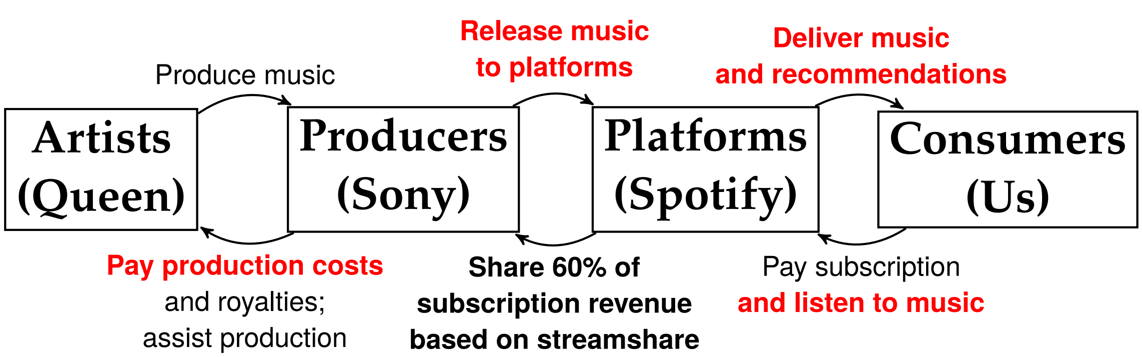

Figure 3: Vertical Structure in the Music Industry

This structure identifies the economics relationships in the music industry, and highlights the ones I estimate in red.

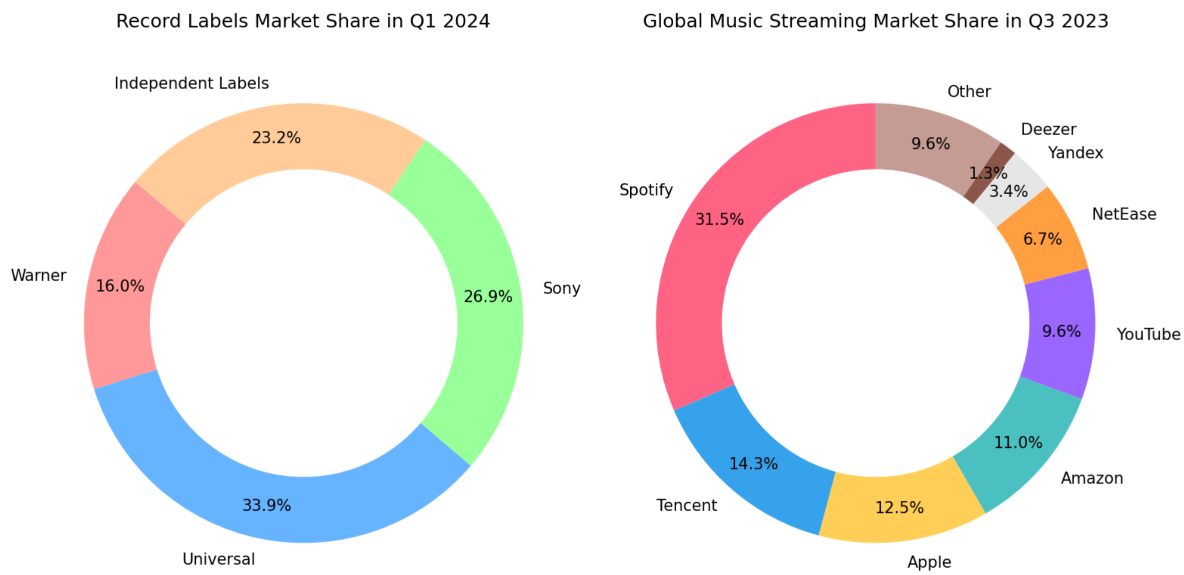

Figure 4: Concentration in the Recorded Music Industry, 2023

The Big Three record labels (WMG, Sony, UMG) comprise 77% of the market, while independent labels comprise the remaining 23%. The Big Three streaming platforms (Spotify, Apple Music, Amazon Music) comprise 80% of the market.

each other, but with the entire history of recorded music. The Bureau of Labor Statistics estimates that there are approximately 35,000 musicians and singers in the U.S., as of May 2023.[13]

Rightsholders, such as Sony, Warner, and Universal (the Big Three record labels), are responsible for distributing music to consumers, either through physical media (e.g., CDs) or through digital platforms (e.g., Spotify). They also search for new and upcoming artists to sign to contracts and promote their music. These labels also have a wide variety of subsidiary labels (or sub-labels) to focus on particular types of music or audiences. These sublabels sometimes end up competing for artists. Rightsholders also negotiate with streaming platforms to distribute music, bargaining over the share of revenue they receive from the platform, and the terms of the contract. I discuss the bargaining between rightsholders and streaming platforms in more detail in the following subsection. Rightsholders are a highly concentrated section of the industry, with the Big Three (WMG, Sony, and UMG, including their sublabels) capturing 77% of the market. Other independent labels comprise the remaining 23% of the market. Figure 4 shows the market share of rightsholders (and streaming services).

Streaming platforms, such as Spotify, Apple Music, and Amazon Music, are respon-

13. https://www.bls.gov/oes/current/oes272042.htm

sible for distributing music to consumers, either through a subscription or ad-supported model. These platforms began to enter the U.S. market in the early 2010s, after starting in Europe in the late 2000s, and they have revolutionized the recorded music industry, allowing consumers to access a vast catalog of music for a fixed monthly fee. As with rightsholders, this section of the industry is highly concentrated, with five firms comprising approximately 80% of the market. Figure 4 shows the market share of streaming platforms (and rightsholders).

These platforms have relatively homogeneous music catalogs, hosting songs from the Big Three and many independent labels. Instead, they differentiate instead on their recommendation engines, interface, and ancillary features (e.g., exclusive podcasts, integration with smart devices, etc.). I speculate that the presence of YouTube as a free, ad-supported platform for music and lyric videos made it difficult for these platforms to compete on exclusive content.[14] This is especially true because non-rightsholders can easily upload music to YouTube, creating a difficult cat-and-mouse game for uploaders, rightsholders and the platform. It is easier for rightsholders to upload their music to YouTube and gain ad revenue for it, thereby making YouTube a streamer of last resort for consumers.

Streaming platforms offer multiple services to consumers, which I group into two: ad-supported and premium subscriptions. Ad-supported subscriptions allows consumers to access music at no monetary cost, instead facing use restrictions and advertising. On Spotify, ad-supported consumers have total access to fifteen playlists, which are a mixture of editorial (human-curated) and algorithmically-generated playlists. For any other playlist on the service, users can only shuffle songs (i.e., they cannot directly select a song). Additionally, ad-supported users can only skip up to six songs per hour, must listen to advertising breaks during their streaming sessions, and stream at lower audio quality (bitrate). Premium subscribers pay a monthly fee ($11.99 a month at the time of writing, $9.99 at the time of analysis) to remove all the aforementioned restrictions.[15]

2.2.1 Vertical Contracts between Rightsholders and Streaming Platforms

Spotify contracts with rightsholders to distribute music to consumers. These contracts set the terms under which Spotify can license music and how Spotify pays rightsholders.[16] Spotify pays rightsholders for royalty-bearing streams (RBS), defined as any play of a

14. While some music platforms (e.g., TIDAL) attempted to differentiate through exclusive music, they abandoned this strategy.

15. Spotify also offers a variety of group and student subscriptions at a lower price per user.

16. Singleton (2015)

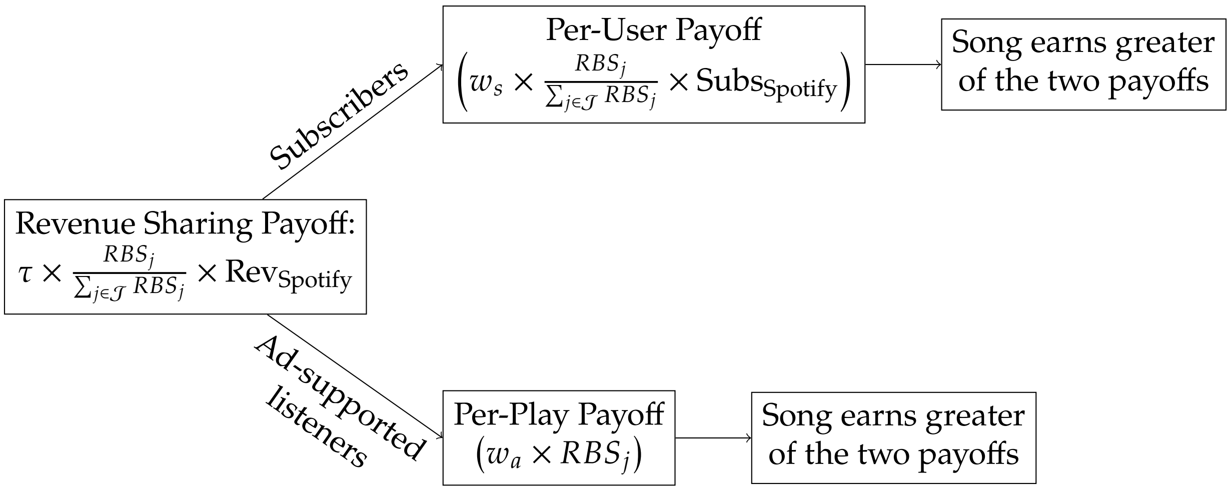

Figure 5: Revenue Sharing Payoff Structure

Rightsholders usually receive a percentage of Spotify's revenue proportionate to their streamshare, but can receive a per-user/per-play fee as a fallback for premium subscribers and ad-supported listeners, respectively.

song that lasts more than 30 seconds.[17] Rightsholders earn income based on their song's streamshare, which is its number of royalty-bearing streams divided by the total number of royalty-bearing streams on the platform in a given month. I write the streamshare equation as follows:

$$\text{Streamshare}_j = \frac{\text{RBS}_j}{\sum_k \text{RBS}_k}$$

Spotify pays rightsholders separately for ad-supported and subscription consumers, and these two types of consumers have different payment structures. For premium subscribers, Spotify pays rightsholders the greatest of a share of gross revenue or a per-subscriber fee, multiplied by a sharing parameter. For ad-supported subscribers, Spotify pays rightsholders the greatest of a share of ad revenue of a per-stream fee. Figure 5 shows the payoff structure for rightsholders.

At the time Spotify entered the market in 2011, its contract with Sony stated that the revenue share was 60%, the per-subscriber fee was \$6, and the per-stream fee was \$0.0225. The contract also had a most-favored nation clause, suggesting that these rates prevailed for all three of the major labels. Spotify has since renegotiated these rates, but the exact terms are not public.

At launch, Spotify charged \$9.99 for a premium subscription, so the revenue share and per-subscriber fee were equivalent at that time. Since Spotify has gone public in 2017, its

17. Spotify has begun to deploy longer cutoffs for certain types of songs to qualify for RBS. https://artists.spotify.com/en/blog/modernizing-our-royalty-system

premium average revenue per user has been well below the per-subscriber fee, primarily because of family and student plans, which reduce the price per user. Assuming that Spotify has not renegotiated the per-subscriber fee with rightsholders, this would suggest that this fee (times the number of subscribers) is greater than the revenue share, and that Spotify is paying rightsholders the per-subscriber fee. Singer and Rosenblatt (2023) suggest, however, that the per-subscriber fee is a floor, and that Spotify pays rightsholders a revenue share of approximately 65% of gross revenue.[18]

The structure of this contract is vital for understanding the incentives of rightsholders to release different kinds of music on Spotify. Firms have a clear incentive to reduce song length to increase the number of RBS, and thereby increase their streamshare and revenue from Spotify. Spotify, however, would pay more for ad-supported subscribers if more streams occurred, so they would prefer to have longer songs. Consumers also have preferences over song length, which can affect these incentives.

Spotify responds to these incentives through its recommender system. Singer and Rosenblatt (2023) report that Spotify's recommender system rewards songs that users complete, and penalize ones that consumer only partially listen to. This has driven rightsholders to adjust the structure and characteristics of their music to align with the priorities of Spotify's recommender system. I investigate how rightsholders have responded to the recommender systems, and whether these recommender systems are welfare-improving.

3 Data

I leverage two sources of data in this project: the Music Streaming Sessions Dataset (MSSD, Brost, Mehrotra, and Jehan 2018), and data from Spotify Charts. The MSSD consists of 160m consumer-level streaming sessions between July 15th and September 18th of 2018, with each session containing up to twenty songs a consumer interacted with on Spotify. The MSSD defines a streaming session as any listening session with less than 60 seconds between songs. The data also only contain streaming sessions with at least ten songs, and it truncates all streaming sessions after twenty songs.

The MSSD contains both song characteristics for the approximately 3 million songs in its data and data for each of the approximately 2bn song-consumer interactions. The song characteristics include both musical characteristics and machine learning characteristics. Musical characteristics include tempo, duration, key, time signature, and mode. Machine learning characteristics are data generated by machine learning classification systems, and these characteristics include danceability, energy, valence, and acousticness. Machine

18. Specifically, labels receive 52%, and publishers receive another 10-12%.

learning characteristics are continuous on a $[0, 1]$ support, while musical characteristics may be continuous (e.g., tempo) or discrete (e.g., key).

Consumer-song interactions include a wide array of information about the consumer and how they interact with the song. The variable of interest is how long the consumer listens to the song, which is grouped into four bins ("skipped very early","skipped early", "listened to most of the song", "listened to the entire song"). I assume that consumers who do not skip a song very early (i.e., are not in the first bin) have listened to enough of the song for it as an RBS. I also observe details about the consumer's streaming session: the position of the song in the session, the date and hour when they listened to each song in the session, and whether the consumer was listening to a song they searched for, their own collection, an editorial playlist, or an algorithmic playlist or radio station. Additionally, I observe what the consumer did after each song, which I use to determine under what circumstances a consumer ended their streaming session. Moreover, I observe the consumer's subscription status at the time of listening. I use these choice-level data to estimate my model of consumer demand and the recommender system.

When working with the MSSD, I use a stratified sampling strategy. Specifically, I sample 0.5% of the consumers who listen to each song. For each song, I sample all of that consumer's streaming session. Additionally, the same consumer may be sampled in multiple songs, but I only include their data once. This results with a sample of 180m observations, representing approximately 10% of the total data.

My second data source is Spotify Charts, a website that reports the top 200 songs on Spotify daily for each country Spotify operates in. For each of these top 200 songs, Spotify reports the number of streams, providing market-level consumption information for these top 200 songs. Spotify also provides the song's Spotify ID, which can be connected to Spotify's API to retrieve the song's characteristics. I rely on a Kaggle dataset that scraped Spotify Charts and Spotify's API to collect this data.[19] I use these data, in conjunction with the demand and recommender system estimates, to estimate the supply model of the industry and to conduct counterfactual analysis.

Another data source to which I have access is the LFM-2B. This dataset contains 2bn listening events from Last.FM, a music scrobbling service. Users can connect their listening histories to Last.FM, which records them and provides recommendations and analysis of their listening habits. These data are available through a public API, and they have been consolidated into a single dataset by Melchiorre et al. (2021). These data contain all listening events from 2005 to 2020, including the song, how long a user listened, and some demographic information about the user: age, gender, country. ListenBrainz is a similar

19. https://www.kaggle.com/edumucelli/spotifys-worldwide-daily-song-ranking

	Mean	Median	Standard Deviation	Min	Max
Duration (s)	203.27	199.32	54.28	30.13	943.53
Release Year	2019	2019	1.39	2017	2021
Acousticness	0.23	0.13	0.25	0.00	0.99
Danceability	0.67	0.68	0.15	0.06	0.98
Energy	0.62	0.63	0.17	0.01	1.00
Instrumentalness	0.01	0.00	0.09	0.00	0.96
Liveness	0.18	0.13	0.14	0.02	0.97
Loudness	-6.83	-6.38	2.71	-38.86	0.35
Mode	0.61	1.00	0.49	0.00	1.00
Speechiness	0.15	0.09	0.13	0.02	0.97
Tempo (BPM)	122.41	122.08	30.04	40.32	212.06
Time Signature	0.97	1.00	0.16	0.00	1.00
Valence	0.46	0.46	0.22	0.03	0.98

Table 3: Spotify Charts Song Characteristics ($N = 9,244$ songs)

Songs in Spotify's Top 200 between 2017 and 2021 were predominantly from those years, high in energy and danceability, and low in acousticness and speechiness.

service, which has become more popular in recent years, and provides similar information as the LFM-2B. I plan to use these data to augment my demand estimates, and to provide more comprehensive listening histories to improve the recommender system model.

3.1 Descriptive Statistics

Table 3 reports the descriptive statistics for the Spotify Charts data.

I focus on the top 200 songs in the US between 2017 and 2021. In this period, $9,244$ unique songs entered Spotify's top 200. The average song length is 3 minutes and 24 seconds, with a standard deviation of 54 seconds. However, the range of length is very wide, with songs as short as 30 seconds and as long as 15 minutes and 45 seconds making it to the top 200. The average song tempo is 122 beats per minute (BPM), with a low of 40 BPM and a high of 212 BPM. All the machine learning characteristics are bounded between 0 and 1, but their averages vary widely: the average song has an average danceability of 0.67, but an average acousticness of 0.23. The average song is an uptempo, energetic, and danceable track, unlikely to be a live recording or acoustic performance. It's also unlikely to be a spoken word song, but it could convey either positive or negative emotion (the valence is 0.46).

Figure 6 reports the correlation matrix of the song characteristics in the Spotify Charts data.

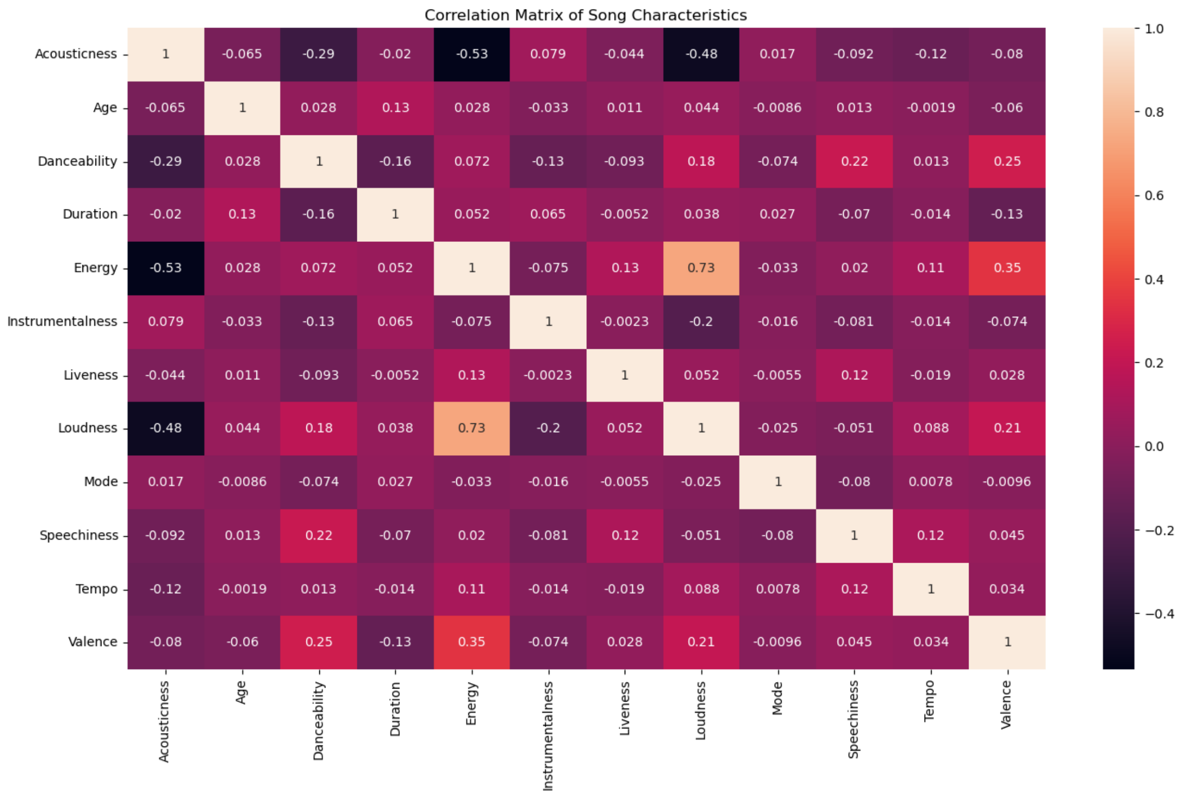

Figure 6: Correlation Matrix of Spotify Charts Song Characteristics

Most of the song characteristics are uncorrelated with each other.

Most of these characteristics are uncorrelated with each other, except for loudness and energy, which are positively correlated (0.73), and loudness and acousticness, which are negatively correlated (-0.53).

The Spotify Charts data also provides information about the lifecycle of songs. Figure 7 reports the number of streams of a song by day after release:

This figure shows the average number of streams each song that made it on Spotify's Top 200 received in the days since its release. Unsurprisingly, songs get a significant number of their streams in the first 100 days after release, with the average number of streams above $400,000$ per day for the first 100 days. After that, the number of streams decreases, with a small uptick around the one and two-year marks, but continuing to fall off over time. The number of streams becomes more volatile after the three-year mark, because fewer songs have been out for that long in my data.

Moreover, I plot the network of songs to determine how much external validity analysis of the US data provides. Figure 8 shows the network of songs in the Spotify Charts data.

Each node (circle) in the chart represents a country, and each edge (line) represents

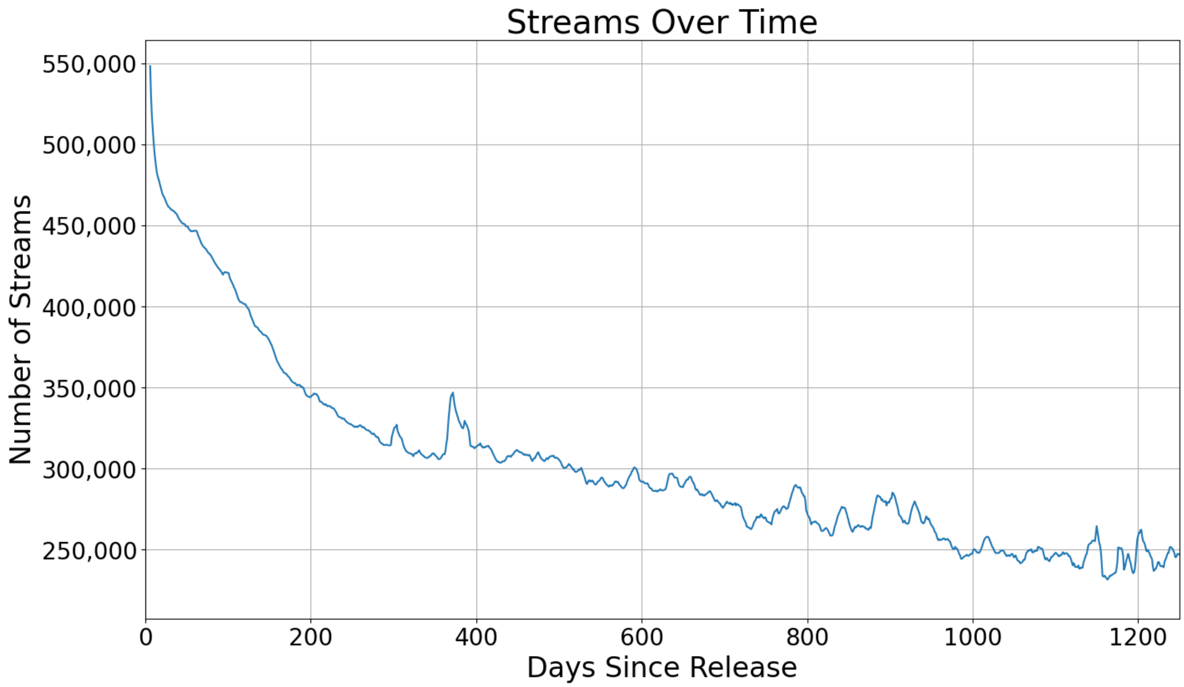

Figure 7: Number of Streams of Songs on Spotify's Top 200, by Days since Release

Songs receive most of their streams in the first 100 days after release, with a small uptick around the one and two-year marks.

songs that appears in the top 200 in both countries. I use a nearest neighbor algorithm to determine which countries have the most overlap with up to 15 neighbors. I then group them by similarity and plot the network. The network has two main clusters: Spanish-speaking countries, and the rest of the world. The rest of the world is highly connected, with significant overlaps in songs. Within the rest of world cluster, some subclusters are apparent: Nordic countries, East Asian countries, and Anglophone countries. This network suggests that focusing on the US provides a good level of external validity for other non-Spanish speaking countries, but that the Spanish-speaking countries may have different preferences in music.

Table 4 reports the descriptive statistics for the songs in the Music Streaming Sessions Dataset.

The MSSD contains approximately 3.7 million unique songs, with an average length of 3 minutes and 54 seconds, with a standard deviation of 1 minute and 48 seconds. Compared to the Spotify Charts data, these songs are longer and have a higher standard deviation in length. These songs are also older than the Spotify Charts songs, with an average release year of 2009 (median 2013), compared to 2019 (median 2019) for the Spotify Charts songs. The songs in these data have similar tempos and levels of energy and valence, but vary

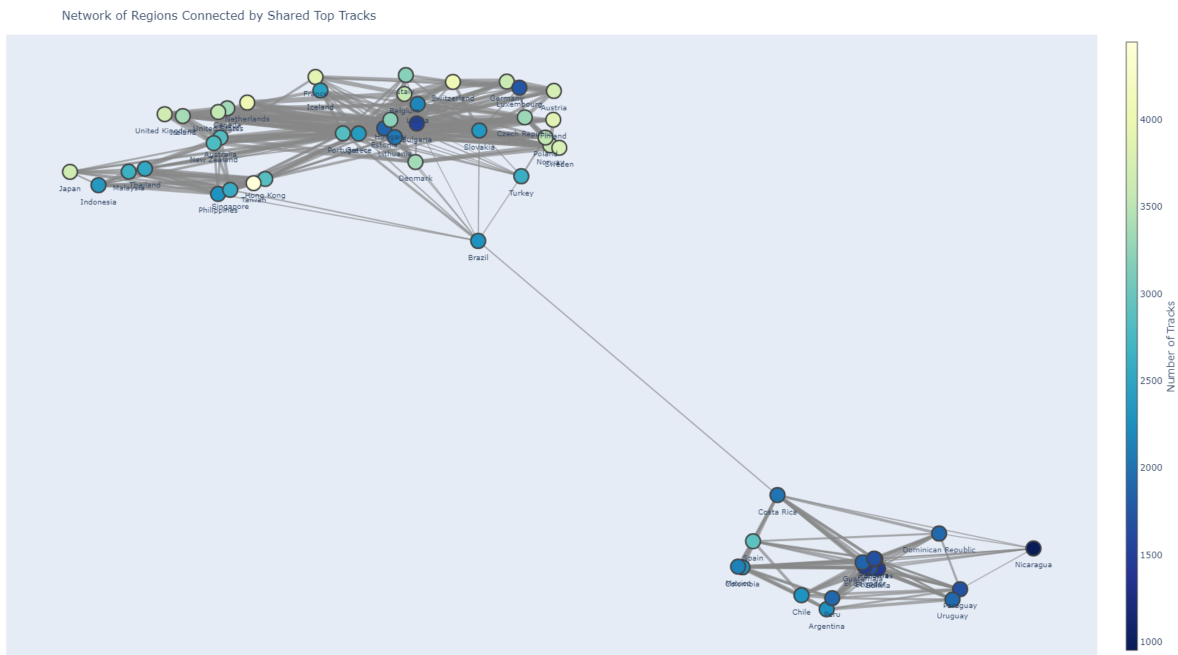

Figure 8: Spotify Charts Song Network

The network of songs in the Spotify Charts data shows two main clusters: Spanish-speaking countries, and the rest of the world.

	Mean	Median	Standard Deviation	Min	Max
Duration (s)	233.19	217.91	108.40	30.00	1800.00
Release Year	2009	2013	11.03	1950	2019
Acousticness	0.35	0.22	0.34	0.00	1.00
Danceability	0.56	0.57	0.19	0.00	1.00
Energy	0.59	0.63	0.26	0.00	1.00
Instrumentalness	0.21	0.00	0.34	0.00	1.00
Liveness	0.21	0.13	0.19	0.00	1.00
Loudness	-9.60	-8.08	5.73	-60.00	6.28
Mode	0.65	1.00	0.48	0.00	1.00
Speechiness	0.10	0.05	0.14	0.00	0.97
Tempo (BPM)	120.07	119.95	30.43	0.00	249.99
Time Signature	0.97	1.00	0.18	0.00	1.00
Valence	0.48	0.47	0.27	0.00	1.00

Table 4: MSSD Song Characteristics ($N = 3.7m$ songs)

Songs in the MSSD are longer, older, and more varied in their characteristics than the Spotify Charts data.

slightly in other characteristics, such as danceability and instrumentalness. Overall, the difference in the data is representative of the changes in popular music over the last decade,

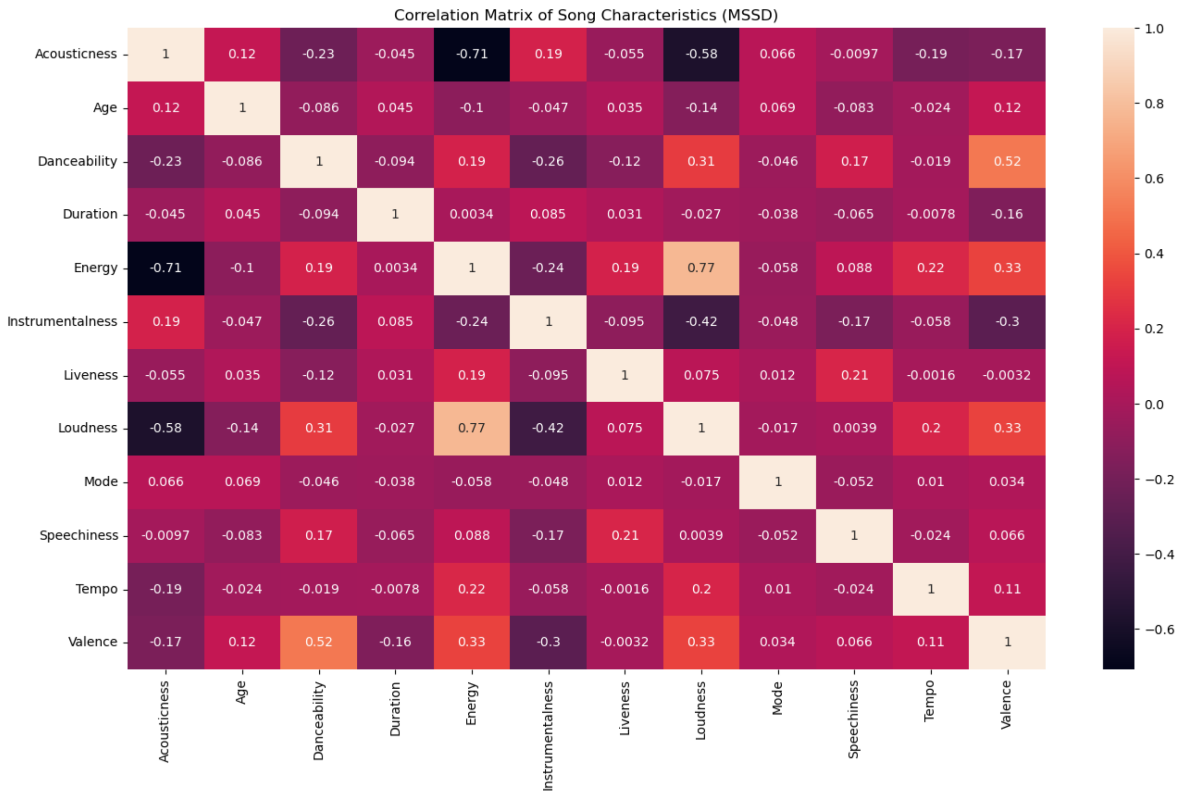

Figure 9: Correlation Matrix of MSSD Song Characteristics

Most of the song characteristics are uncorrelated with each other.

with the MSSD data representing a wider variety of music than the Spotify Charts data. Specifically, the Spotify Charts data reflects more spoken-word, danceable, and shorter songs. When using both of these datasets, I standardize the Spotify Charts variables using the MSSD variables.

Figure 9 reports the correlation matrix of the song characteristics in the MSSD.

As with the Charts data, most of these characteristics are uncorrelated, but with the same exceptions: loudness and energy are positively correlated (0.77), and loudness and acousticness are negatively correlated (-0.58). Energy and acousticness are also negatively correlated (-0.71), as valence and danceability are positively correlated (0.52).

Table 5 reports the consumer-level statistics for my sample of the Music Streaming Sessions Dataset.

Consumers in my sample are primarily premium subscribers, with 84% of the sample being premium subscribers. This is significantly higher than the percentage of premium subscribers Spotify reports, which is 40% of its user base.[20] It is, however, more repre-

20. Spotify Q2 2024 Earnings Report

	Mean	Standard Deviation
Session Length	18.07	6.91
% Shuffle	0.35	0.31
% Premium Subscribers	0.84	0.31
% RBS	0.58	0.31
% Completion	0.34	0.31
% Morning Listen	0.24	0.31
% Afternoon Listen	0.39	0.31
% Evening Listen	0.29	0.31
% Night Listen	0.08	0.27
% Monday Listen	0.15	0.31
% Tuesday Listen	0.15	0.31
% Wednesday Listen	0.14	0.31
% Thursday Listen	0.14	0.31
% Friday Listen	0.15	0.31
% Saturday Listen	0.13	0.31
% Sunday Listen	0.13	0.31
% Catalog Listen	0.24	0.43
% Chart Listen	0.01	0.11
% Editorial Playlist Listen	0.15	0.35
% Algorithmic Playlist Listen	0.03	0.16
% Algorithmic Radio Listen	0.15	0.35
% User Collection Listen	0.42	0.49

Table 5: MSSD Consumer Characteristics ($N = 180m$ song-consumer interactions)

Consumers in the MSSD have long streaming sessions, with a high percentage of RBS, but a low percentage of song completion. They primarily listen to their own collections, but about 20% of their listens are algorithmically driven.

sentative of the percentage of revenue Spotify earns from premium subscribers, which is 88% of its revenue.[21] These users have very active streaming sessions, with an average session length of 18 songs. They also are somewhat likely to listen on shuffle, with 35% of sessions being shuffle sessions. These listeners are also rather active: while 58% of consumer-song interactions are long enough to be considered an RBS, consumers only complete 34% of the songs they receive. Listening time is even throughout the week, with 13-15% of sessions occurring on each day of the week. Within a day, however, very little listening occurs at night (12-6 AM), with only 8% of sessions occurring during this time.

Consumers in my sample primarily listen to music from their own search process, or

21. Spotify Q2 2024 Earnings Report

from their own collections, with 66% of sessions being from these sources. Algorithmic playlists and radio stations consist of 18% of streaming sessions. Editorial (human-curated) playlists and top charts are the least common source of music, with only 16% of sessions coming from these playlists.

4 Model

To model the effect of recommender systems on the music industry, I develop a structural model of the industry, with three sets of agents: consumers, a recommender system, and rightsholders. Consumers (the demand side) receive songs from the platform (and its recommender system) and choose whether to listen to them. I capture this choice using a random utility model, which generates a probability of listening to a song based on its characteristics and the consumer's characteristics. The recommender system, which I treat as an exogenous technology, computes the probability consumers receive particular songs based on their characteristics and the consumer's characteristics. The recommender system surfaces songs in proportion to their probability of being listened, and the joint probability of being surfaced and the probability of being heard is the choice probability rightsholders face. On the supply side, rightsholders choose whether to release songs provided to them by artists, paying a fixed cost to releasing them. Rightsholders (the supply side) choose whether to release the song they have in their inventory, based on its expected profit, which is a function of the choice probabilities at the time of release and in the future. These rightsholders are forward-looking, anticipating the evolution of the market and the recommender system through first-order Markov processes. To motivate these processes, I employ and oblivious equilibrium, where each firm considers only the long-run average choice of the industry, rather than each rival's choice. Figure 10 describes the timing of the model each period.

4.1 Demand

Consumers in my demand model are subscribers to a streaming platform offering them a catalog of songs.[22] Each day, these consumers open the streaming app and start receiving songs from the platform, as informed by the recommender system. For each song they receive, consumers make one of three possible choices: listen to the song (up to the amount necessary for an RBS), skip the song, or stop listening to the platform, which I treat as an

22. I do not model the extensive decision to subscribe to Spotify (or join the ad-supported tier). While Spotify does report subscriber data, price variation is somewhat limited over time.

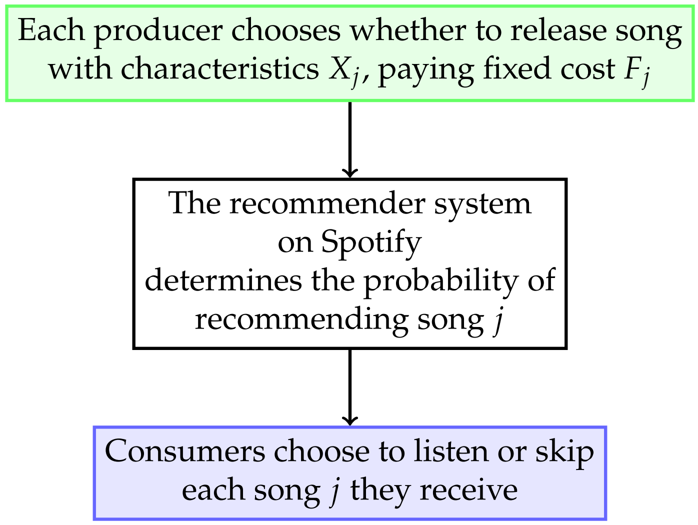

Figure 10: Timing of the Model in Each Period

Producers move first in the model, followed by the recommender system, then consumers. I solve this model recursively.

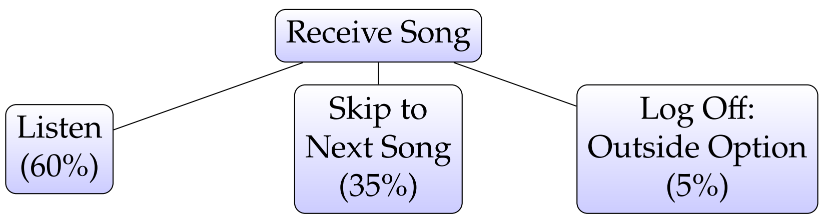

Figure 11: Consumer Decision Tree

Consumers, after receiving a song, choose whether to listen to a song, skip it, or log off, ending their streaming session.

outside option. Figure 11 describes the decision tree for consumers in the demand model.

I maintain one assumption about consumers in my model:

Assumption 1 *Consumers do not consider how their choice affects future personalized recommendations.*[23]

This assumption allows me to model consumers as static agents, simplifying the demand model and allowing me to focus on the supply-side effects more directly.

Consumers have random utility over the songs they receive and the outside option. Consumer i's utility of listening to a particular song j in session position s is given by:

23. Anecdotal evidence suggests consumers do not extensively think about future songs when deciding whether to listen to a song, or how their choice affects future recommendations, especially when they are uninformed about the specific mechanisms of the recommender system.

$$U_{L,ijs} = \beta X_j + \gamma_L Y_i + \eta_{Ls} + \epsilon_{ijs} \tag{2}$$

In this utility function, X_j are a vector of linear and quadratic song characteristics (alternative-specific variables), Y_i are a vector of consumer characteristics (case-specific variables), η_{Ls} are position-specific fixed-effects, and ϵ_{ijs} is a Type 1 (Gumbel) Extreme Value error term. The consumer characteristics and session fixed effects are case-specific variables, so their parameters are also case-specific, following the standard in conditional choice models. Intuitively, consumers prefer certain types of music, which I decompose into quantitative characteristics, and their utility from a particular song may depend on when they are listening, both during the day, and where they are in their streaming session. Additionally, to capture horizontal preferences over music, I employ quadratic terms for the song characteristics, which allow for non-linear preferences. Passive consumers may not skip songs often (if at all); active users are likely to skip songs often, finding one they like; and hybrid consumers may skip early in the streaming session before settling on a set of songs they enjoy, and listening to them.

I normalize the mean utility of the outside option to zero:

$$U_{i0s} = \epsilon_{i0s} \tag{3}$$

4.1.1 Utility of Skipping Songs

To capture the utility of skipping to the next song, consumers form expectations over the characteristics of the next song, based, generally, on the songs they have received in their streaming session so far. Their utility from skipping has the following equation:

$$U_{S,ijs} = \beta E_{is}[X_j | X_{j,s-1}] + \gamma_S Y_i + \eta_{Ls} + \epsilon_{ijs} \tag{4}$$

I refine these expectations using the listening context data from the MSSD. Specifically, I apply the following rules:

- If consumers are listening to an algorithmic playlist or radio station, then their expected utility of skipping comes from the average characteristics of the songs they have received in their streaming session so far.

- If consumers are listening to their own catalog or playlist, or a song they searched for, then their expected utility of skipping comes from the average characteristics of the songs in their entire streaming session.

- If consumers are listening to editorial playlists or top 200 playlists, then their expected utility of skipping depends on whether they shuffle the playlist: if they do, expected utility comes from the characteristics of songs received so far; if not, then the expected utility comes from the average characteristics of the songs in streaming session.

Intuitively, consumers know more about their own playlists, music catalog, or searches, so their expectations will be more refined than just discovering music on an algorithmic playlist. If they are listening to an editorial playlist or top 200 playlists, I use shuffling as a proxy for awareness of songs on the playlist: consumers who do not shuffle may be more aware of the tracks on the playlist, and therefore more aware of their characteristics, than those who do not.

4.1.2 Choice Probabilities

In this model, consumers choose whether to listen to the song they receive, to skip it, or to log off, ending their streaming session and taking an outside option.

The T1EV error term in the utility function allows me to model the choice probabilities as a conditional logit model. The probability that consumer i listens to song j in session position s, conditional on the song being recommended, is given by:

$$\begin{aligned} &P(i \text{ listens to } j | \text{RS surfaces } j \text{ to } i) \\ &= \frac{\exp(\beta X_j + \gamma Y_i + \eta_s)}{1 + (\exp(\beta X_j + \gamma Y_i + \eta_s) + \exp(\beta E_{is}[X_j | X_{j,s-1}] + \gamma Y_i + \eta_s))} \end{aligned} \tag{5}$$

4.2 Recommender System

Recommender systems are an integral component to music streaming, directing consumers towards songs the system thinks they will enjoy. These recommender systems are functionally trying to solve a multi-armed bandit problem: finding the best product (arm) to offer to consumers (slot machines), with success being a purchase or interaction with the product. To train the optimal recommender system, platforms must balance exploration (trying new products) and exploitation (recommending products that are likely to be successful). Firms typically rely on an ϵ-greedy algorithm, where the firm chooses the best product with probability $1 - \epsilon$, and a random product with probability ϵ.

I group these systems into three types: collaborative filtering recommender systems, content-based recommender systems, and hybrid recommender systems. Collaborative filtering recommender systems surface products based on products similar users like. For example, if person 1 likes songs, X, Y, and Z, and person 2 likes songs W, X, and Y, then the system may recommend song Z to person 2 and song W to person 1.[24] Content-based recommender systems decompose products into characteristics, and recommend products with similar characteristics to those the user has liked in the past. For example, if person 1 likes songs with a high tempo, the system may recommend songs with a high tempo to person 1.[25] Hybrid recommender systems combine aspects of both collaborative filtering and content based recommender systems. Most recommender systems are hybrid, albeit weighted towards one end or the other.

Spotify's recommender system is a hybrid system weighted heavily towards content-based recommendations. They use a combination of user and song characteristics to recommend songs to users. While the recommender system itself is a black box, various research papers have discussed its mechanisms, and I use these papers for guidance in constructing my model of the recommender system, particularly McInerney et al. (2018).

McInerney et al. (2018) describes Spotify's recommender system as having an objective (or reward) function with the following form:

$$r_{ij} = \sigma(\iota_1 X_j + \iota_2 Y_i)$$

In this equation, r_{ij} is the binary outcome from recommending a song j to listener i. X_j are the song characteristics, and Y_i are the listener characteristics. ι_1 and ι_2 are the parameters to be trained. σ is a sigmoid loss, making this equation a logistic regression. McInerney et al. (2018) further augment this function with higher-order interactions between the user and consumer characteristics to obtain more personalized recommendations. They also interact these terms to further personalize the recommendations. To implement the recommender system, they use a standard ϵ-greedy algorithm.

I use a logistic regression to model Spotify's primarily content-based recommender system. I treat this recommender system as an exogenous technology to which Spotify has access, and I estimate the parameters of the recommender system using data from the MSSD. I assume for simplicity that, when Spotify is recommending songs, they are following a pure exploitation strategy, rather than an ϵ-greedy strategy.

I estimate the recommender system using the following equation:

$$P(\text{RS surfaces } j \text{ to } i) = \frac{\exp(\eta_1 X_{1j} + \eta_2 \prod_{n=2}^{N} X_{nj} + \eta_3 Y_i)}{1 + \exp(\eta_1 X_{1j} + \eta_2 \prod_{n=2}^{N} X_{nj} + \eta_3 Y_i)} \tag{6}$$

24. Amazon uses collaborative filtering when recommending products "people like you also bought".
25. Continuing the Amazon example, they use content-based recommendations when describing "similar products".

Here, $P(\text{RS surfaces } j \text{ to } i)$ is estimated probability that Spotify recommends song j to consumer i. X_{1j} are song characteristics from music theory, and X_{nj} are machine learning characteristics, interacted with each other. Y_i are consumer characteristics, and η_1, η_2, and η_3 are parameters to be estimated. Unlike in my choice model, the outcome variable $P(\text{RS surfaces } j \text{ to } i)$ is a listen to completion, rather than just enough to qualify as an RBS. The recommender system also places no value on skipping a song, whereas consumers may have some expected utility for skipping a song (e.g., to find a song they like more). I take equation 6 to the MSSD data.

Having described the recommender system and the choice model, I combine these two models to create the demand producers face:

$$\begin{aligned} P(i \text{ listens to } j) = {} & P(\text{RS surfaces } j \text{ to } i) \times P(i \text{ listens to } j | \text{RS surfaces } j \text{ to } i) \\ & + (1 - P(\text{RS surfaces } j \text{ to } i)) \times P(i \text{ listens to } j | \text{RS does not surface } j \text{ to } i) \end{aligned} \tag{7}$$

For producers, consumers can access their songs in two ways: through the recommender system, or through other means. I treat these other means as the complement to the probability the recommender system surfaces the song. Intuitively, it can also represent a function representing consumer awareness of the song outside the recommender system. Because I observe whether consumers are receiving songs from the recommender system or from other sources, I estimate separate parameters for each listening type.

This approach builds on Goeree (2008), who using a joint probability to create a demand structure. She uses this structure to model the demand for computers when consumers have limited information. In place of a recommender system, she uses advertising to inform the consumers and construct consideration sets. I do not explicitly construct consideration sets, because my choice structure is a sequence of binomial listen/skip choices (with an outside option), rather than a single multinomial choice.

4.3 Supply

Rightsholders are the supply side of the music industry, choosing whether to release songs to Spotify. They are forward-looking agents, considering both current and future profits when making their decision. Rightsholders face a fixed cost to release a song, and they receive revenue each period based on that song's streamshare.[26]

Each rightsholder receives a song from an artist, knowing its characteristics, and they

26. I treat revenue from Spotify as exogenous, because I do not model Spotify as a strategic agent.

decide whether to pay the fixed cost to release the song on Spotify. In making this decision, rightsholders consider both the probability the recommender system will amplify their song, and the probability consumers will listen to their song. I maintain one assumption about rightsholders in my model:

Assumption 2 *Each song has an independent rightsholder (i.e., no multi-product competition), and each song has an exogenous release date, so firms face a one-time binary release/no-release decision.*

Once a song is on Spotify, it remains on the platform in perpetuity, so rightsholders can earn revenue in future periods. To effectively make this decision, they must have some way to model future period profits. Specifically, rightsholders need to model two sets of evolutionary processes:

- The evolution of rival songs, which affects the probability consumers listen to their song
- The evolution of the recommender system (i.e., the probability their song is recommended to consumers)

I define $\mathcal{X}_t$ as the mean characteristics of all songs on a given day on Spotify Charts, and I define ϕ as the probability the recommender system recommends a song to a consumer in future periods. With these terms defined, I now define the following first-order Markov processes by which the recommender system and rival songs evolve:

$$\mathcal{X}_{t+1} = \nu_0 + \nu_1 \mathcal{X}_t + \epsilon_t^{\mathcal{X}} \tag{8}$$

$$\phi_{j,t+1} = \psi_0 + \psi_1 \phi_{jt} + \epsilon_{jt}^{\phi} \tag{9}$$

To motivate these processes, I use an oblivious equilibrium (Weintraub, Benkard, and Van Roy 2005) as my solution concept. This equilibrium is a typically used to analyze dynamic oligopoly models with a large number of firms. In an oblivious equilibrium, firms make decisions based only on their own state and average industry conditions, ignoring the specific states of their competitors. Weintraub, Benkard, and Van Roy (2005) show that, under certain conditions, the oblivious equilibrium is equivalent to the Markov Perfect Nash Equilibrium. This simplification allows me to tractably estimate my supply model while capturing the key dynamics in the industry.

As applied to my model, each firm is an oblivious agent, choosing whether to release its song based on their song's characteristics, the long-run average characteristics of all songs, and the probability the recommender system will recommend their songs. Recall that each song has its own rightsholder, so each song competes with every other song in the market, past, present, and future, resulting in thousands of firms.

Having explained how rightsholders act in the model, as well as the motivating solution concept, I now define their expected profit function:

$$E[\pi_j(X_j)] = 0.6 \left(\sum_{t=0}^{T} \delta^t R_t \left(\frac{P(i \text{ listens to } j \text{ with characteristics } X)}{\sum_K P(i \text{ listens to } k \text{ with characteristics } \mathcal{X})} \right) \right) - F_j \tag{10}$$

Each period t, defined as a day, the rightsholder owning song j receive a share of Spotify's gross revenue R_t. I define this share as follows:

$$s_{jt} = \frac{R\hat{B}S_{jt}}{\sum_k R\hat{B}S_{kt}} = \frac{P(i \text{ listens to } j \text{ with characteristics } X)}{\sum_K P(i \text{ listens to } k \text{ with characteristics } \mathcal{X})}$$

This share is the streamshare of song j in period t. δ is the firm's discount factor. F_t is the fixed cost to release song j on Spotify, which varies by day.

This share can be further simplified:

$$\begin{aligned} s_{jt} &= \frac{R\hat{B}S_{jt}}{\sum_k R\hat{B}S_{kt}} \\ &= \frac{P(\text{RS surfaces } j \text{ to } i) \times \exp\left(\beta X_j + \gamma_L Y_i + \eta_{Ls}\right)}{\sum_k P(\text{RS surfaces } j \text{ to } i) \times \exp\left(\beta X_k + \gamma_L Y_i + \eta_{Ls}\right)} \\ &+ \frac{1 - P(\text{RS surfaces } j \text{ to } i) \times \exp\left(\beta X_j + \gamma_S Y_i + \eta_{Ss}\right)}{\sum_k 1 - P(\text{RS surfaces } j \text{ to } i) \times \exp\left(\beta X_k + \gamma_S Y_i + \eta_{Ss}\right)} \end{aligned}$$

Having decomposed equation 10, we now turn to the entry condition. Because each rightsholder faces a one-time binary decision to release or not release, they release as long as the following condition holds:

$$0.6 \left(\sum_{t=0}^{T} \delta^t R_t \left(\frac{P(i \text{ listens to } j \text{ with characteristics } X)}{\sum_K P(i \text{ listens to } k \text{ with characteristics } \mathcal{X})} \right) \right) \geq F_j \tag{11}$$

Specifically, their expected profit from releasing the song must be nonnegative. If the expected revenue exceeds fixed cost, the rightsholder releases the song; otherwise, it does not.

This entry condition provides the upper bound to the fixed cost of releasing a song. I use data on gross profit margins to scale the expected revenue to a fixed cost observation.

I assume that fixed costs are independently and identically distributed along a lognormal distribution. I augment this with a heirarchical model, where the parameters of the lognormal distribution are functions of the song characteristics. Formally, I define F_j as follows:

$$F_j \sim lognormal(\mu_j(X_j), \sigma_j(X_j))$$
$$\mu_j(X_j) = \mu_g + X_j' \beta_\mu$$
$$\sigma_j(X_j) = \sigma_g + X_j' \beta_\sigma$$

μ_g and σ_g are global parameters representing the population-level mean and standard deviation. β_μ and β_σ are k-dimensional vectors of coefficients capturing the effects of covariates on the mean and variance, respectively.

4.4 Equilibrium

My solution concept is an oblivious equilibrium (Weintraub, Benkard, and Van Roy 2005) where consumers optimally choose whether to listen or skip songs in their streaming session, or to log off; the recommender system optimally recommends songs to consumers, seeking to maximize the probability consumers listen to songs to completion; rightsholders, taking the above as given, choose whether to release songs based on the expected profit from releasing the song; songs enter the market as long as their expected revenue is greater than their fixed cost.

5 Estimation

My estimation strategy has several stages:

1. Demand and Recommender System estimation

2. Markov Process estimation

3. Expected revenue calculation

4. Fixed cost estimation

In the first stage, I estimate consumer preferences and recommender system preferences using the MSSD data. I construct separate estimates for consumers who receive songs from the recommender system, and those who do not. Specifically, I estimate

$\theta_1 = (\beta, \gamma, \eta)$ from equations 5 and 6 in this stage. For consumer preferences, I use a maximum likelihood estimator over choice probabilities, following Train (2009). I identify my parameters through variation in the choices each consumer faces at each position in the streaming session. Similarly, I estimate the recommender system parameters using a maximum likelihood estimator over the probability a consumer completes a song.

In the second stage, I estimate the Markov processes governing the evolution of right-sholder perception of the recommender system and rival songs. Specifically, I estimate $\theta_2 = (\nu_0, \nu_1, \psi_0, \psi_1)$ in this stage. To construct the Markov process for the recommender system, I use θ_1 to predict the probability the recommender system will surface a song to a consumer, and I compute the average of these probabilities across all songs in the Top 200 each day. I then estimate a SARIMAX model for ψ_0 and ψ_1. For the song characteristics, I compute the average characteristics of all songs on Spotify's Top 200 each day, and I estimate the ν_0 and ν_1 as a Vector Autoregression (VAR) model.

In the third stage, I compute the expected revenue for each song released in 2018. I limit my computation to songs released between January 1, 2018, and September 30, 2018, to better match my demand and recommender system estimates. For each song, I compute the left-hand term in equation 10, using the θ_1 and θ_2 estimates to predict future streamshare. In the specification with consumer awareness, I conservatively assume that consumers become aware (outside the recommender system) of a song one year after release.

When computing the expected revenue, I assume that consumers are premium subscribers, and that they listen in the evening. These are the modal consumer characteristics in the MSSD data. I also assume that the song is the first one in their streaming session to normalize the expected revenue estimates.

To compute the rival songs in the streamshare measure, I take $\mathcal{X}$ for the songs available on the top 200 in the day the song has been released, and I input these characteristics to predict the probability the recommender system will surface the rival song. I then apply that predicted probability to the VAR(1) process to estimate the probability the recommender system will surface the rival song in future periods. I also apply these characteristics to the VAR(1) process to estimate the probability consumers will listen to the rival song.

To account for only observing the top 200 songs in constructing the streamshare measure, I estimate the total number of streams on Spotify in a given day, and then downscale the amount of revenue to match the percentage of streams coming from the top 200 songs. I first assume that the average listener on Spotify spends 125 minutes listening to music each day.[27] Next, I take the average length of streams of Spotify's Top 200 songs to estimate the number of songs a listener listens to each day. I then multiply this by the reported number of users on Spotify to estimate the total number of streams on Spotify each day. Finally, I divide the number of streams of the top 200 songs by the total number of streams to estimate the percentage of streams coming from the top 200 songs, and I use this percentage to downscale the revenue to match the revenue generated by the top 200 songs.

As an alternative approach, I assume that the number of rival songs (each possessing the same characteristics) is equal to the number of songs on the platform, which is approximately 40m in 2018.[28] This creates a lower bound for the amount of revenue any given song can generate, but coheres with the idea that each song competes with every other song on the platform.

Having computed expected revenue, I now use these data as the input to the likelihood function for the lognormal distribution. I first scale the expected revenue by gross profit margins observed in the music industry, by label, and treat these scaled revenues as the fixed costs observations from the lognormal distribution.[29] I estimate the fixed cost parameters using maximum likelihood estimation. The log-likelihood function for a single observation is:

$$\ell_j(\theta) = -\frac{1}{2}\log(2\pi\sigma_j^2) - \frac{(\log(F_j) - \mu_j)^2}{2\sigma_j^2}$$

where $\theta_3 = (\mu_g, \sigma_g, \beta_\mu, \beta_\sigma)$ is the vector of parameters to be estimated. The full log-likelihood is:

$$L(\theta) = \sum_{i=1}^{n} \ell_j(\theta) \tag{12}$$

I solve 12 for θ_3 using a BFGS algorithm with numerical gradients.

6 Results

Tables 6 and 7 reports consumer demand estimates for both direct selection and recommender system selection:

27. I take an average of the reported listening time of the following two industry reports: IFPI and Global Web Insights

28. Spotify 2018 Annual Report

29. I use the Earnings before Interest, Taxes, Depreciation, and Amortization (EBITDA) margin for the major labels to scale the expected revenue for their songs, and a 20% EBITDA margin for independent labels.

	Dependent variable: Royalty Bearing Stream (RBS)			
	Direct Selection		Recommender System Selection	
	Estimate (Robust Std. Error)	Odds Ratio	Estimate (Robust Std. Error)	Odds Ratio
Age	-0.0112***	0.999	0.00161	1.000
	(0.000492)		(0.00168)	
Acousticness	0.0604***	0.661	0.0423***	0.689
	(0.000639)		(0.00196)	
Acousticness^2	-0.0456***	–	-0.0378***	–
	(0.000111)		(0.000252)	
Danceability	-0.0402***	0.777	0.150***	3.287
	(0.00112)		(0.00362)	
Danceability^2	-0.0112***	–	0.0651***	–
	(0.000475)		(0.00124)	
Duration	0.00342***	1.000	-0.00254	1.000
	(0.000839)		(0.00296)	
Duration^2	-0.000722***	–	0.00759***	–
	(0.0000857)		(0.000446)	
Energy	-0.00595***	1.605	-0.0648***	4.552
	(0.00132)		(0.00437)	
Energy^2	0.0459***	–	0.161***	–
	(0.000398)		(0.00106)	
Instrumentalness	0.00115**	0.738	-0.0368***	0.672
	(0.000523)		(0.00158)	
Instrumentalness^2	-0.0101***	–	-0.00628***	–
	(0.0000479)		(0.000111)	
Liveness	0.0160***	0.811	0.00530***	1.061
	(0.000533)		(0.00177)	
Liveness^2	-0.00832***	–	0.000740***	–
	(0.0000677)		(0.000168)	
Loudness	0.000418	1.000	0.0488***	1.010
	(0.00103)		(0.00355)	
Loudness^2	-0.00600***	–	-0.0136***	–
	(0.000103)		(0.000361)	
Mode	0.0308***	1.031	0.0484***	1.050
	(0.001)		(0.00313)	
Speechiness	0.00132**	0.540	-0.00170	0.629
	(0.000511)		(0.00170)	
Speechiness^2	-0.0115***	–	-0.00829***	–
	(0.0000612)		(0.000149)	
Tempo	-0.00257***	1.000	0.0184***	1.000
	(0.000935)		(0.00289)	
Tempo^2	-0.0149***	–	0.00976***	–
	(0.000482)		(0.00113)	
Time Signature	0.0415***	1.042	-0.0597***	0.942
	(0.00318)		(0.0109)	
Valence	0.0256***	0.497	0.0155***	0.995
	(0.000749)		(0.00234)	
Valence^2	-0.0730***	–	-0.00517***	–
	(0.000263)		(0.00059)	
Model Statistics				
Observations	148,822,923		31,238,428	
$\bar{\rho}^2$	0.018		0.284	

Notes:
*** p<0.01, ** p<0.05, * p<0.1

Table 6: Consumer Demand Estimates - Song Characteristics

Demand within and outside the recommender system is similar, except for song energy, danceability, and valence.

	Dependent variable: Royalty Bearing Stream (RBS)			
	Direct Selection		Recommender System Selection	
	Estimate (Robust Std. Error)	Odds Ratio	Estimate (Robust Std. Error)	Odds Ratio
Time of Day				
Morning	0.131***	1.140	0.150***	1.162
	(0.00136)		(0.00415)	
Afternoon	0.102***	1.107	0.121***	1.129
	(0.00121)		(0.00371)	
Night	0.0417***	1.043	0.125***	1.133
	(0.00191)		(0.00611)	
Day of Week				
Tuesday	0.0108***	1.011	0.0289***	1.029
	(0.00180)		(0.00538)	
Wednesday	0.0163***	1.016	0.0231***	1.023
	(0.00184)		(0.00552)	
Thursday	0.00907***	1.009	0.0207***	1.021
	(0.00183)		(0.00551)	
Friday	0.00165	1.002	-0.0520***	0.949
	(0.00179)		(0.00529)	
Saturday	-0.00358*	0.996	-0.0220***	0.978
	(0.00186)		(0.00568)	
Sunday	-0.00389**	0.996	-0.0117**	0.988
	(0.00184)		(0.00566)	
User Characteristics				
Premium	-0.121***	0.886	0.116***	1.123
	(0.0014)		(0.00465)	
Model Statistics				
Observations	148,822,923		31,238,428	
$\bar{\rho}^2$	0.018		0.284	

Notes:
*** p<0.01, ** p<0.05, * p<0.1

Table 7: Consumer Demand Estimates - Contextual Characteristics

Demand within and outside the recommender system is similar, except for premium subscription status.

The results reveal striking differences between how users interact with music through direct selection versus recommender systems. In direct selection, song characteristics have more modest effects, with most musical features showing relatively small coefficients. Age has a small negative effect on direct demand, suggesting users slightly prefer newer songs when choosing directly. Interestingly, acousticness and valence (emotional positivity) show positive linear terms but negative quadratic terms in direct selection, indicating users prefer moderate levels of these characteristics when actively choosing music.

The recommender system demand parameters show notably larger magnitudes and sometimes different directions of effects compared to direct selection. Most strikingly, danceability and energy have much stronger positive effects in recommended songs (with odds ratios of 3.287 and 4.552 respectively) compared to their modest or negative effects in direct demand. This suggests the recommender system may be effectively surfacing high-energy, danceable songs that users might not have chosen directly but end up enjoying. The model fit $\bar{\rho}^2$ is also substantially better for recommender system demand (0.284) compared to direct demand (0.018), indicating that song characteristics are more predictive of consumption when songs are recommended.

The contextual characteristics also reveal interesting patterns. Premium users are less likely to engage with directly chosen songs (odds ratio 0.886) but more likely to engage with recommended songs (odds ratio 1.123), suggesting they may be more trusting of or receptive to recommendations. They might also be more passive listeners, preferring to let the recommender system guide their listening. Time of day effects are stronger for both types of demand during morning and afternoon hours, with slightly larger coefficients for recommended songs. Day of week effects show that both types of demand are higher during midweek and lower on weekends, though the negative weekend effect is stronger for recommended songs, particularly on Fridays (odds ratio 0.949).

Tables 8 and 9 reports the results for the recommender system:[30]

The recommender system's song selection behavior differs notably from how users engage with its recommendations. While both models show that age negatively influences recommendations and consumption, the effect is much stronger in the demand model (coefficient -0.0432 vs 0.00161), suggesting the system may be too aggressive in favoring newer songs. Similarly, for characteristics like speechiness and instrumentalness, the system shows strong positive preferences (odds ratios of 1.618 and 1.419 respectively) that aren't matched in user engagement, where these features actually show negative coefficients in the demand model. This mismatch suggests potential areas where the recommender system could be better aligned with user preferences.

30. Introducing interactions does not materially affect many of these estimates.

	Dependent variable: Song Completion	
	Estimate (Std. Error)	Odds Ratio
Age	-0.0432***	0.995
	(0.000168)	
Acousticness	0.0395***	1.058
	(0.000280)	
Acousticness²	-0.0065***	–
	(0.000189)	
Danceability	-0.0530***	1.021
	(0.000414)	
Danceability²	0.0181***	–
	(0.000439)	
Duration	-0.2395***	0.998
	(0.000322)	
Duration²	0.0161***	–
	(0.0000999)	
Energy	0.0178***	1.345
	(0.000454)	
Energy²	0.0222***	–
	(0.000383)	
Instrumentalness	0.0304***	1.419
	(0.000661)	
Instrumentalness²	0.0059***	–
	(0.000162)	
Liveness	-0.0307***	1.192
	(0.000286)	
Liveness²	0.0098***	–
	(0.000106)	
Loudness	-0.0235***	0.995
	(0.000430)	
Loudness²	0.0022***	–
	(0.0000996)	
Mode	0.0113***	1.011
	(0.000329)	
Speechiness	-0.0815***	1.618
	(0.000267)	
Speechiness²	0.0199***	–
	(0.0000847)	
Tempo	-0.0076***	1.000
	(0.000308)	
Tempo²	0.0212***	–
	(0.000503)	
Time Signature	0.0336***	1.116
	(0.001081)	
Valence	0.0345***	0.943
	(0.000254)	
Valence²	-0.0160***	–
	(0.000260)	
Model Statistics		
Observations	180,061,351	
Pseudo R^2	0.006733	

Notes:
*** p<0.01, ** p<0.05, * p<0.1

Table 8: Recommender System Estimates - Song Characteristics

The recommender system prioritizes shorter, more energetic songs with a standard time signature

Dependent variable: Song Completion		
	Estimate (Std. Error)	Odds Ratio
Time of Day		
Morning	0.1531***	1.165
	(0.000437)	
Afternoon	0.0963***	1.101
	(0.000392)	
Night	0.1706***	1.186
	(0.000624)	
Day of Week		
Tuesday	-0.0082***	0.992
	(0.000570)	
Wednesday	-0.0088***	0.991
	(0.000584)	
Thursday	-0.0177***	0.982
	(0.000583)	
Friday	-0.0215***	0.979
	(0.000571)	
Saturday	-0.0166***	0.984
	(0.000597)	
Sunday	-0.0061***	0.994
	(0.000591)	
User Characteristics		
Premium	-0.0673***	0.935
	(0.000434)	
Model Statistics		
Observations	180,061,351	
Pseudo R^2	0.006733	
Notes:		
*** p<0.01, ** p<0.05, * p<0.1		

Table 9: Recommender System Estimates - Contextual Characteristics

Premium consumers are less likely to complete songs

The energy and danceability parameters reveal particularly interesting dynamics. While the demand model showed users strongly engage with high-energy, danceable recommended songs (odds ratios of 4.552 and 3.287), the recommender system model shows more modest positive effects for energy (odds ratio 1.345) and even a slight negative coefficient for danceability. This suggests the system may be under-recommending songs with these characteristics relative to user preferences. The quadratic terms for many characteristics also differ between the models, indicating that the system's understanding of optimal levels for these features might not perfectly align with what drives user engagement.

The contextual effects show some alignment but also key differences. The time-of-day patterns are similar, with both demand and consumption showing higher activity during morning and night hours. However, the day-of-week patterns differ notably: while the demand model showed higher engagement during midweek, the recommender system model shows consistently negative coefficients for all days relative to Monday. For premium users, both models show negative effects (odds ratios of 0.935 for provision and 0.886 for demand), suggesting the system may be appropriately calibrated in how it treats premium status. Intuitively, premium subscribers, facing no ad interruptions, may be more likely to skip songs and search, whereas ad-supported users would prefer to avoid ads, and take a more passive approach to listening. The much lower pseudo-R^2 for the recommender system model (0.007 vs 0.284) suggests that these observable characteristics explain far less of the system's recommendations than they do user engagement.

Table 10 reports the results for the song characteristic Markov processes:

This VAR suggests strong, stationary processes for each song characteristic with respect to its own lag. All own-lag coefficients are statistically significant, and all of them are less than 0.95. The drift terms are statistically significant, but they are all very close to zero, further suggesting that the processes are stationary. The constant terms are sometimes significant, and most of the cross-characteristic lags are statistically insignificant. This suggests that the processes are relatively independent of each other.

Table 11 reports the results for the recommender system Markov process estimation:

This SARIMAX model suggests that the recommender system is relatively stable, with a high persistence term, but not so high as to suggest that the system is nonstationary. The drift term is statistically significant, but close to zero, further suggesting that the system is stationary.

Figure 12 plots the distribution of expected revenue for songs released in 2018 that entered Spotify's top 200 at least once:

These songs have an expected revenue ranging from \$20,000 to \$500,000, with a mean

Statistic	Acousticness	Age	Danceability	Duration	Energy	Instrumentalness	Liveness	Loudness	Mode	Speechiness	Tempo	Time Signature	Valence
Constant	0.447**	-0.325	-0.320*	0.028	-0.320**	-0.131***	-0.104	-0.015	0.061	-0.044	-0.122	0.158***	-0.190*
	(0.144)	(0.265)	(0.127)	(0.075)	(0.111)	(0.040)	(0.064)	(0.122)	(0.040)	(0.095)	(0.069)	(0.014)	(0.086)
Drift	0.000**	0.000	-0.000***	-0.000***	-0.000	0.000***	0.000***	-0.000	0.000*	-0.000***	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Own Lag	0.927***	0.832***	0.784***	0.907***	0.858***	0.815***	0.870***	0.900***	0.918***	0.887***	0.860***	0.841***	0.924***
	(0.027)	(0.028)	(0.023)	(0.011)	(0.028)	(0.015)	(0.014)	(0.032)	(0.011)	(0.017)	(0.012)	(0.013)	(0.012)
Observations	1820	1820	1820	1820	1820	1820	1820	1820	1820	1820	1820	1820	1820
AIC	-100	-100	-100	-100	-100	-100	-100	-100	-100	-100	-100	-100	-100
BIC	-99	-99	-99	-99	-99	-99	-99	-99	-99	-99	-99	-99	-99

Note: Standard errors in parentheses. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

Table 10: VAR Estimation - Song Characteristics

The VAR(1) process suggests strong, stationary processes for each song characteristic with respect to its own lag.

	Dependent variable: Predicted Probability ($\hat{\phi}_t$)
$\hat{\phi}_{t-1}$	0.734***
	(0.019)
Drift	0.000***
	(0.000)
Constant	0.092***
	(0.007)
Observations	1826
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 11: Markov Process Estimation for Recommender System

The recommender system has a high, statistically significant, persistence term, suggesting that it is relatively stable.

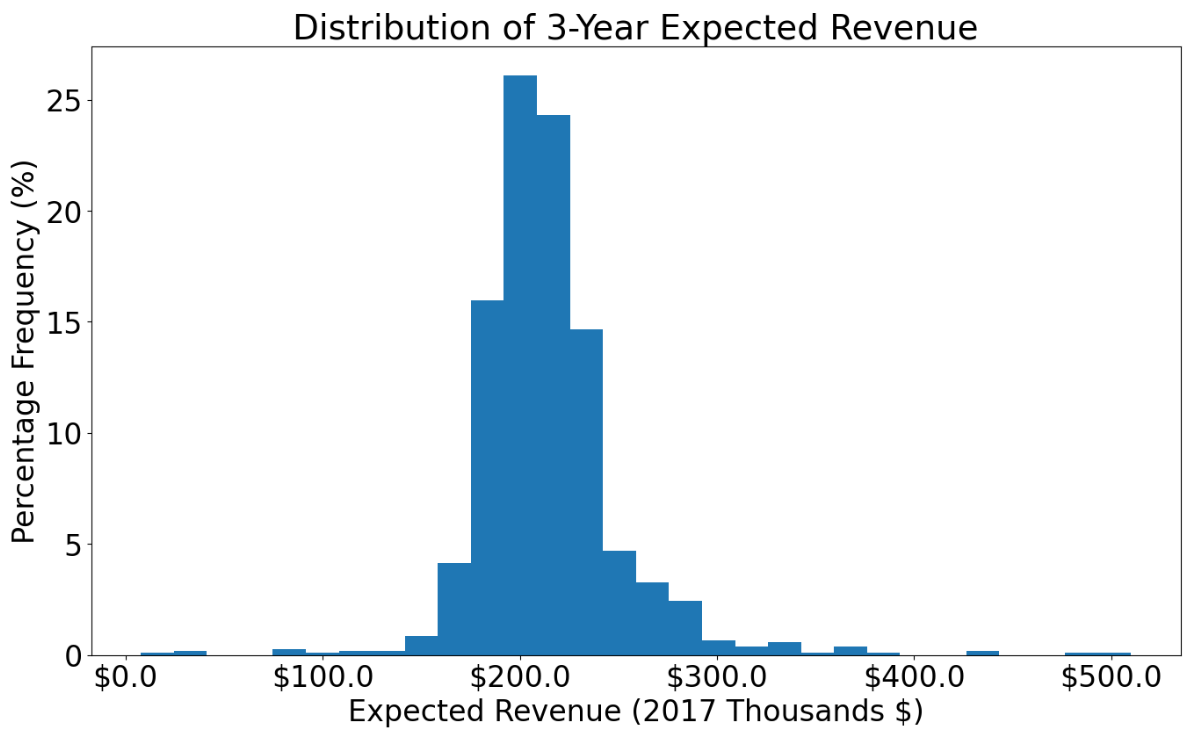

Figure 12: Expected Revenue of Songs Released in 2018 that Entered Spotify's Top 200

These songs have a mean expected revenue of $213,000.

at $213,000. The distribution is similar under my alternative approach, but with much lower amounts, as the rival songs include all songs on the platform. I compute a mean expected revenue of $60.09 for all songs.

My fixed cost estimation produces estimates of the location (μ) and (σ) parameters of

	Location (μ) Estimate (Std. Error)	Scale (σ) Estimate (Std. Error)
Intercept	5.091***	-1.945***
	(0.013)	(0.062)
Audio Features		
Acousticness	0.017***	0.027
	(0.006)	(0.028)
Instrumentalness	-0.056**	0.121**
	(0.022)	(0.060)
Liveness	-0.008	0.153***
	(0.006)	(0.028)
Speechiness	-0.008*	-0.155***
	(0.004)	(0.023)
Loudness	-0.009	-0.272***
	(0.008)	(0.042)
Duration		
Duration	-0.041***	-0.098***
	(0.008)	(0.034)
Duration^2	0.010	0.240***
	(0.008)	(0.033)
Record Label		
Sony	-0.052***	0.101
	(0.015)	(0.086)
Universal	0.062***	-0.075
	(0.012)	(0.068)
Warner	0.038**	0.273***
	(0.017)	(0.087)
Model Statistics		
Observations	1,067	
$\bar{R}^2$	0.108	0.971

Notes:
*** p<0.01, ** p<0.05, * p<0.1

Table 12: Fixed Cost Parameter Estimates

Song characteristics and label fixed effects significantly affect the distribution of fixed costs.

the lognormal fixed cost distribution as a function of song characteristics.

Table 12 reports the results of the fixed cost estimation:

The location model results reveal several key determinants of fixed costs in music production. Duration shows a significant negative linear effect (-0.041), indicating that longer songs generally cost less to produce. Among audio features, instrumentalness has the largest negative effect (-0.056), suggesting that instrumental tracks may be less expensive to produce, possibly due to lower costs associated with not needing vocalists. There are also significant differences across record labels - Universal and Warner songs show higher average fixed costs (coefficients of 0.062 and 0.038 respectively) while Sony songs have lower costs (-0.052), which could reflect different production strategies or accounting practices across labels.

The scale model results reveal interesting patterns in cost variability. Warner shows significantly higher cost variability (0.273), suggesting they may take more risks in production budgets compared to Universal and Sony, whose variance effects are not statistically significant. Among audio features, liveness and instrumentalness are associated with higher cost variability (0.153 and 0.121), while speechiness and loudness are associated with lower variability (-0.155 and -0.272). This suggests that live-recording and instrumental elements introduce more uncertainty into production costs, while speech elements and production choices around loudness may follow more standardized cost structures. Duration shows a U-shaped relationship with cost variability, with the negative linear term (-0.098) and positive quadratic term (0.240) indicating that medium-length songs have the most predictable costs.

I use these estimates to compute the lognormal distribution for each song, and report the median μ for each song as its fixed cost.

Figure 13 plots the distribution of median fixed costs predicted by the model:

The fixed cost of releasing a song on Spotify ranges from \$120,000 to \$220,000, with a median of \$170,000. In comparison, the fixed cost under my alternative approach (competing with all songs) is lower, with a median of \$82.10.

My estimated median fixed cost for songs in the top 200 is close to a report from Chace (2011), which estimated the production and recording costs of Rihanna's "Man Down" at \$78,000 in 2011 dollars (\$88,000 in 2017 dollars). The median for my alternative approach is also close to the fixed cost estimate in Aguiar and Waldfogel (2018). They find that, in their imperfect foresight model, the fixed cost is \$18.97 (\$20.92 in 2017 dollars), approximately \$20 less than my estimate. Several factors explain this difference. First, their model only looks at the revenue generated by the song in 2011. I model songs more dynamically, looking at revenue generated in the first three years of release. Additionally, they estimate a single fixed cost, assuming the fixed cost is the lowest expected revenue for all songs released in a year. In contrast, I estimate a distribution of fixed costs. Moreover, they estimate the fixed cost for a digital release (e.g., on iTunes), which may have different fixed costs than a release on Spotify.

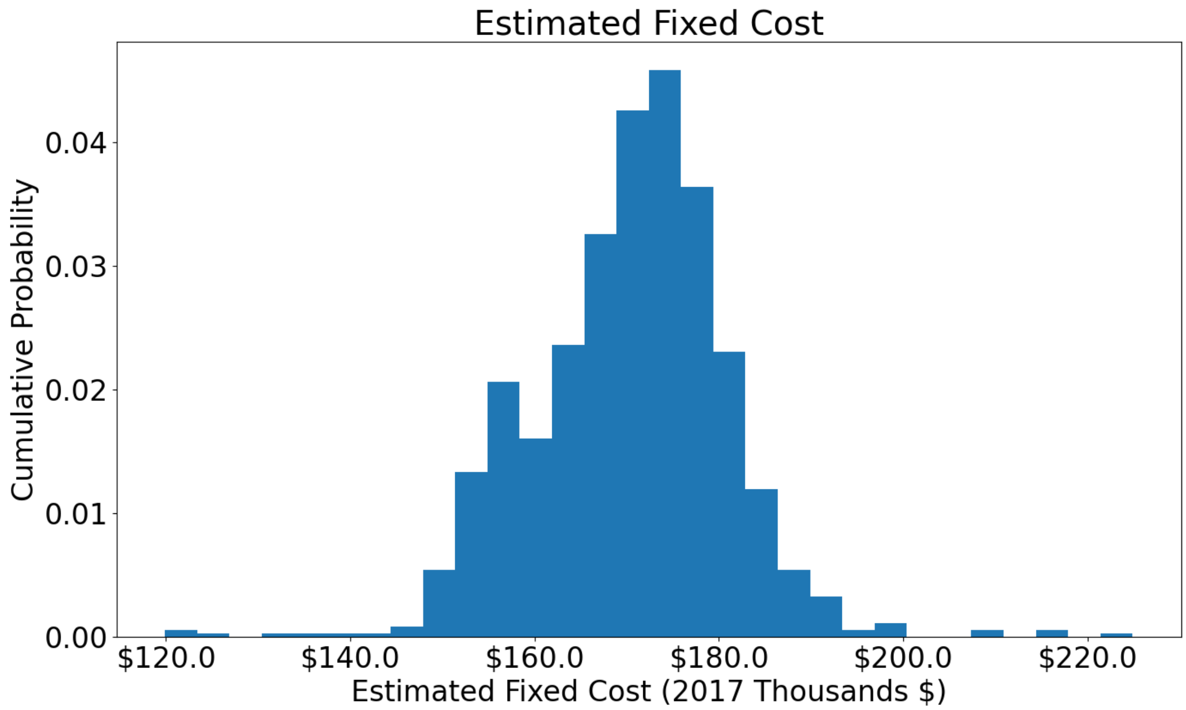

Figure 13: Distribution of Estimated Fixed Costs

These songs have a median fixed cost of $170,000.

7 Counterfactual Analysis

Having estimated demand for song characteristics, the recommender system preferences, and the fixed cost to releasing a song onto Spotify, I now turn to the counterfactual analysis that can answer the question this paper poses: whether recommender systems have affected the kind of music record labels are releasing. To isolate the impact of recommender systems specifically, I conduct two counterfactuals.[31] In the first, I construct a random recommender system, rather than one which relies on song and consumer characteristics.

7.1 Random Recommendations

Intuitively, this random recommender is akin to having no recommender system at all, insofar as the recommendations will be pure noise. It also effectively simulates a naive search process, wherein consumers sample new songs from a uniformly random distribution. I implement this counterfactual by using the following process:

31. I also conduct a simulated counterfactual with an oracular recommender system. See the appendix for details.

1. Draw 500 consumers and give them preferences from the demand estimates. I sample from the distribution of consumer characteristics in the data to construct values for those parameters.

2. Simulate a streaming session of 15 songs for each consumer, drawing from songs released before 2018.

3. Take the average of those songs to generate the utility of skipping a song.

4. Provide a new release to each consumer, and compute the choice probability of listening to the song.

5. Repeat this process for all songs released in the first three quarters of 2018.

6. Compute the expected revenue generated for these new releases, assuming each song has a 25% chance of being recommended, and compare it to the estimated revenue generated by the model.

First, I compare average expected profit for songs. Figure 14 reports the results of this comparison:

Each observation in this figure represents a song released in the first three quarters of 2018. Note that some of the estimated net profits are negative, because the median fixed cost estimated by the parameters of the lognormal distribution is higher than the expected revenue. Intuitively, it is likely that the realized fixed cost for those songs is lower than the median fixed cost. Many songs become unprofitable when random recommendations are used. Indeed, of the 1053 songs I observe that were released in the first three quarters of 2018, 274 (26%) are unprofitable. For those songs that are profitable, their gross profit margin is 11.9%, compared to the average of 20% observed in the industry.

I now turn to some song characteristic results and welfare implications of my counterfactual analysis. Figure 15 reports the average duration of songs between profitable and unprofitable songs.

The average duration of songs of profitable songs is 217 seconds, and the average duration of unprofitable songs is 184 seconds. This difference is significant at the 5% level. Moreover, the unprofitable songs are more homogeneous, as the standard deviation of duration is 1.75 seconds, compared to 2.61 seconds for profitable songs. The difference in distributions is also significant at the 5% level. This suggests that introducing recommender systems allows shorter, more homogeneous songs to enter the market and find an audience.

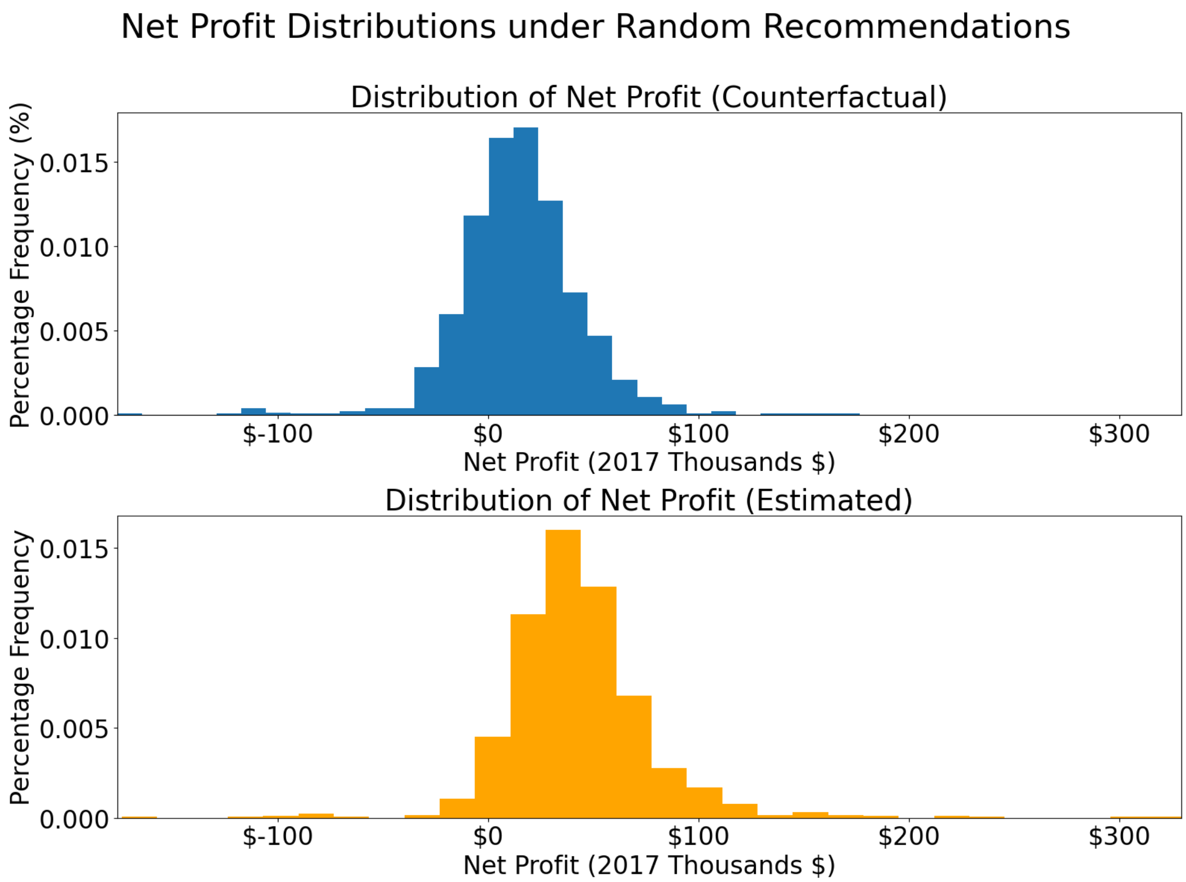

Figure 14: Counterfactual Expected Profit - Random Recommendations

Many songs become unprofitable when random recommendations are used.

Song valence represents another example of the differences between profitable and unprofitable songs. Valence is a measure of the emotional positivity of a song, with higher values indicating more positive emotions. Figure 16 reports the average valence of songs between profitable and unprofitable songs.

The average valence of profitable songs is 0.457, and the average valence of unprofitable songs is 0.385. This difference is significant at the 1% level. Moreover, the distribution of energy is very clearly right-shifted for unprofitable songs, compared to profitable songs.

Table 13 reports the average values of other song characteristics for profitable and unprofitable songs, the difference in means, and the difference in distributions (as evaluated by a KS-Test):

The data reveals several striking differences between profitable and unprofitable songs' musical characteristics under random recommendations, with the Kolmogorov-Smirnov tests indicating significantly different distributions for many key features ($p < 0.001$). Most notably, profitable songs are significantly longer, with an average duration of about

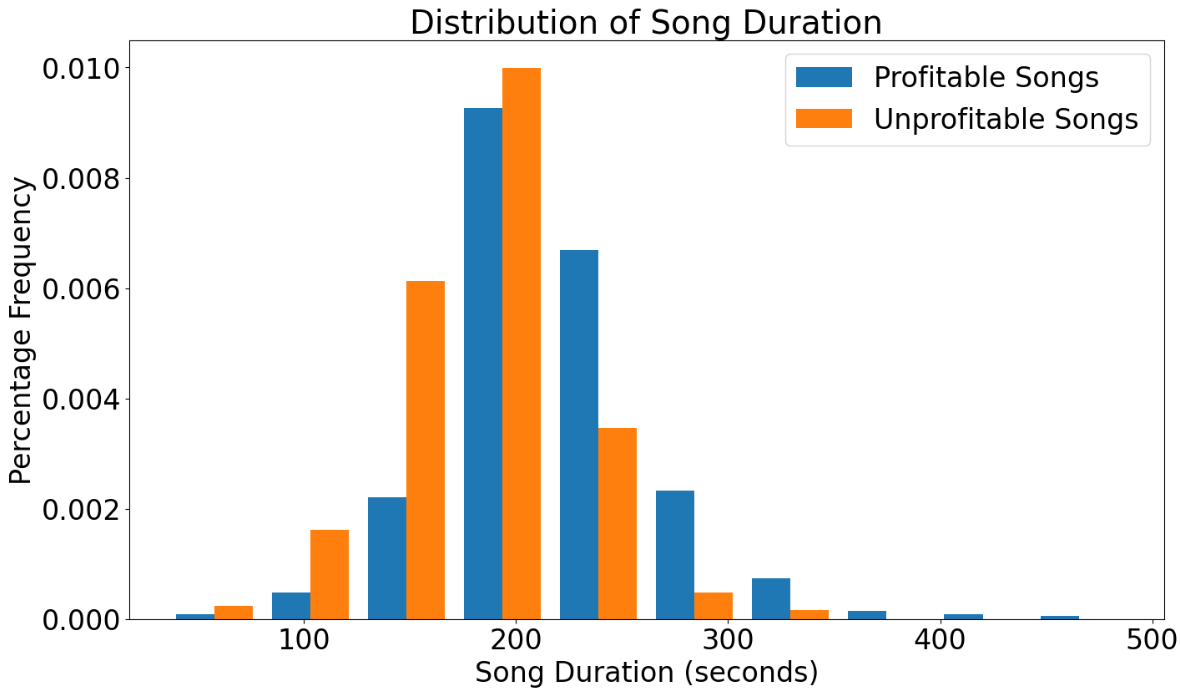

Figure 15: Counterfactual Duration - Random Recommendations

Unprofitable songs are shorter and more homogeneous than profitable songs.

217 seconds compared to 184 seconds for unprofitable songs - a 33-second difference that reflects fundamentally different distributions in song length ($p \approx 0$). Similarly, the distributions of both danceability and valence differ significantly between profitable and unprofitable songs (both $p \approx 0$), with profitable songs showing higher values in both measures (danceability: 0.723 vs 0.619; valence: 0.457 vs 0.385). The mode difference of -0.166 also reflects distinctly different distributions ($p \approx 0$), indicating profitable songs are systematically more likely to be in a major key.

While some characteristics show differences in means, their distributional differences are less pronounced. For instance, despite mean differences in tempo, liveness, and speechiness, their KS test p-values (0.0184, 0.0209, and 0.0212 respectively) suggest more subtle distributional differences that might not be economically meaningful. Most notably, instrumentalness shows virtually identical distributions between profitable and unprofitable songs (KS p-value = 0.9359), despite a small difference in means. Overall, this comparison suggests that the recommender systems allows for shorter, more homogeneous, and more energetic songs to enter the market.

Finally, I turn to the welfare implications of my counterfactual analysis. I compute the consumer surplus generated by all the songs in the release set, as well as the set of surviving songs, by taking the log-sum of the exponentiated utility, following Anderson, Palma, and

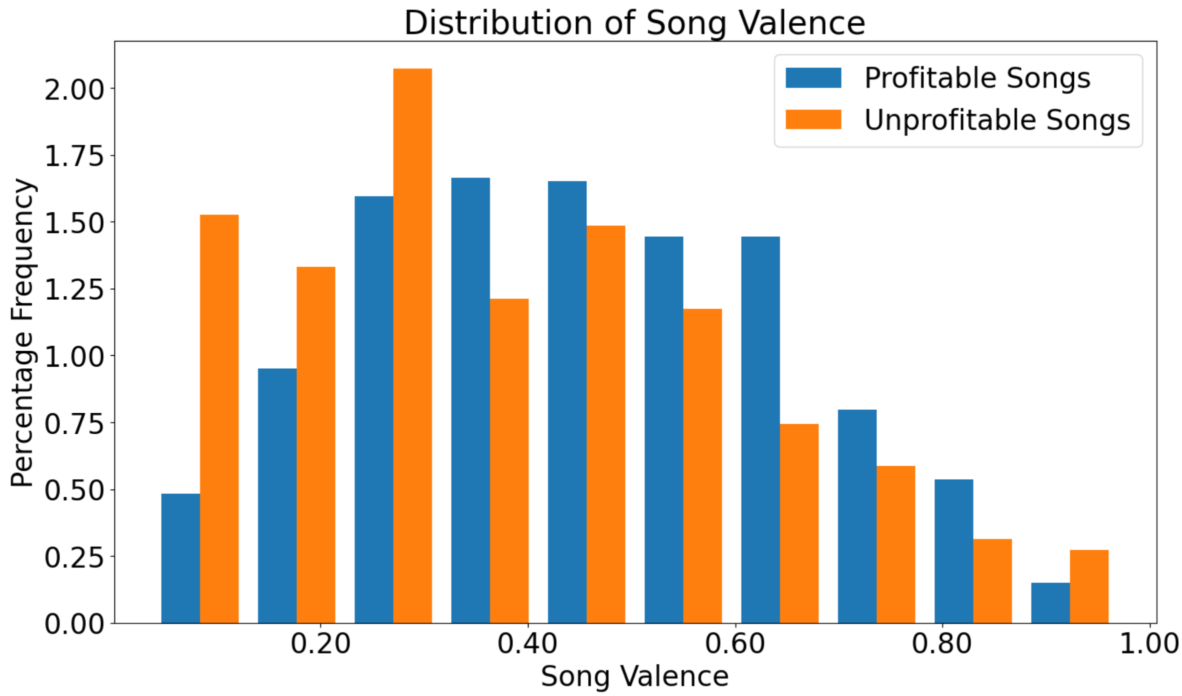

Figure 16: Counterfactual Valence - Random Recommendations

Unprofitable songs are less positive-sounding than profitable songs.

Thisse (1992). Formally, I define consumer surplus with the following equation:

$$CS = \log\left(\sum_{i=1}^{N} \exp\left(\beta X_j + \gamma_L Y_i + \eta_{Ls}\right)\right) \tag{13}$$

Here, N represents the number of songs in the set, rather than the binomial skip-listen decision. Note that this measure of consumer surplus is in utils, as there is no price coefficient against which to scale the results.

I find that consumer surplus is 3.9% higher when targeted recommender systems are used, compared to when random recommendations are used. Restated, random recommender systems result in a 3.8% decrease in consumer surplus. This suggests that recommender systems have increased consumer surplus by allowing for more songs to enter the market, and for consumers to find songs that they enjoy more easily.

7.2 Popular Recommendations

The second counterfactual analysis I conduct is a popular recommender system. It is similar to placing a ban on using consumer data for recommendations, and relying only on the popularity of songs. This recommender system also replicates the market environment

Feature	Profitable	Unprofitable	Diff	KS-Test P-Value
Duration (s)	216.872	183.730	-33.143***	0.0000
	(1.751)	(2.605)		
Tempo	125.093	123.233	-1.860	0.0184
	(1.038)	(2.030)		
Energy	0.628	0.639	0.011	0.0015
	(0.006)	(0.007)		
Danceability	0.723	0.619	-0.104***	0.0000
	(0.005)	(0.007)		
Valence	0.457	0.385	-0.073***	0.0000
	(0.007)	(0.013)		
Acousticness	0.190	0.251	0.061***	0.0004
	(0.007)	(0.017)		
Instrumentalness	0.007	0.012	0.005	0.9359
	(0.002)	(0.005)		
Liveness	0.184	0.167	-0.017	0.0209
	(0.005)	(0.008)		
Speechiness	0.162	0.163	0.001	0.0212
	(0.005)	(0.010)		
Loudness	-6.288	-6.603	-0.315	0.0269
	(0.084)	(0.145)		
Mode	0.641	0.474	-0.166***	0.0000
	(0.017)	(0.030)		

Table 13: Counterfactual Song Characteristics - Random Recommendations

Unprofitable songs are shorter, less danceable, and less positive-sounding than profitable songs.

that existed prior to Spotify, when consumers would purchase singles on iTunes. At the time, the iTunes store did not have a recommender system; instead, it showed users what the top-selling singles and albums were. I replicate this by recommending songs in proportion to their listening shares.

I implement this counterfactual in the following way:

1. Draw 500 consumers and give them preferences from the demand estimates. I sample from the distribution of consumer characteristics in the data to construct values for those parameters.

2. Simulate a streaming session of 15 songs for each consumer, drawing from songs released before 2018.

3. Take the average of those songs to generate the utility of skipping a song.

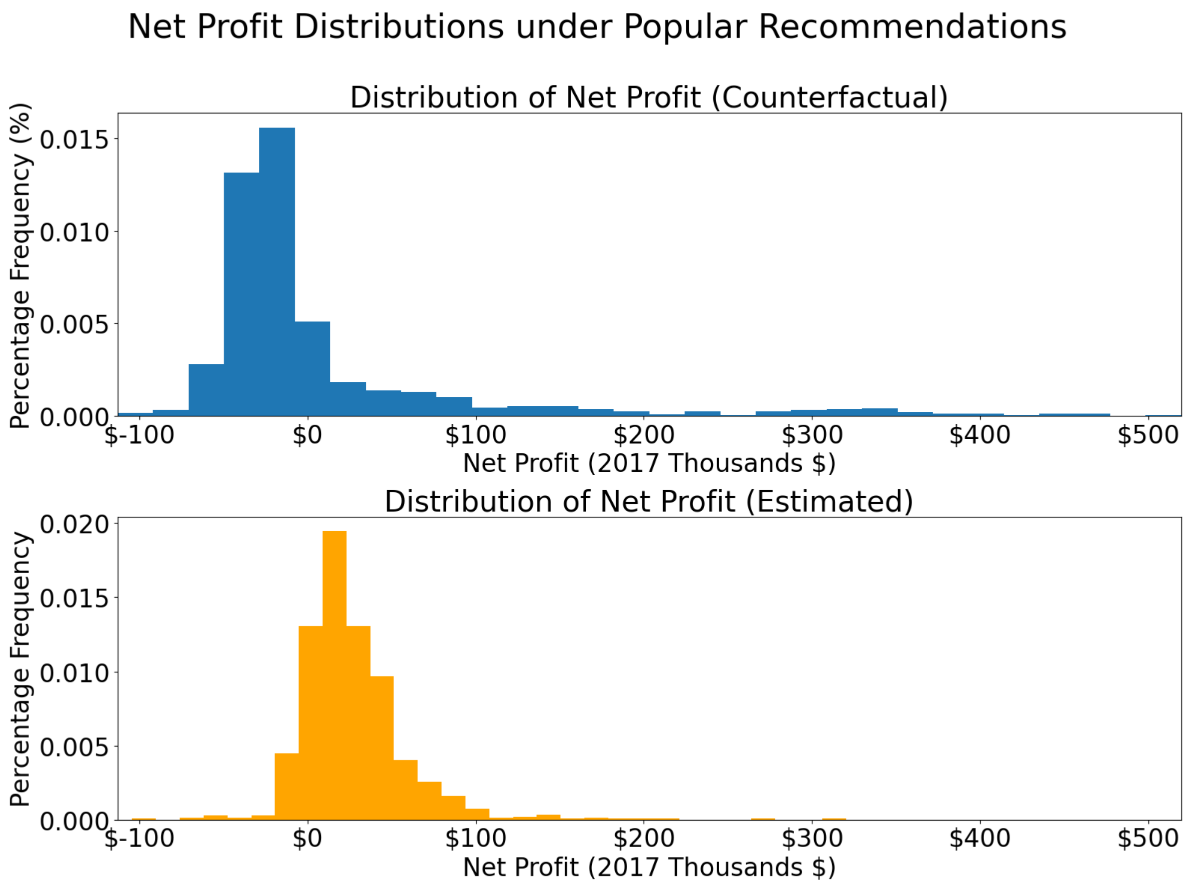

Figure 17: Counterfactual Expected Profit - Popular Recommendations

A superstar effect occurs with popular recommendations: many songs become unprofitable in expectation, but some become highly profitable.

4. Provide a new release to each consumer, and compute the choice probability of listening to the song.

5. Repeat this process for all songs released in the first three quarters of 2018.

6. Compute the share of listens by release day, and set the recommendation probability of each song to be equal to its listening share.

7. Compute the expected revenue generated for these new releases and compare it to the estimated revenue generated by the model.

First, I compare average expected profit for songs. Figure 17 reports the results of this comparison:

Each observation in this figure represents a song released in the first three quarters of 2018. Whereas random recommendations reduced the expected profit of all songs,

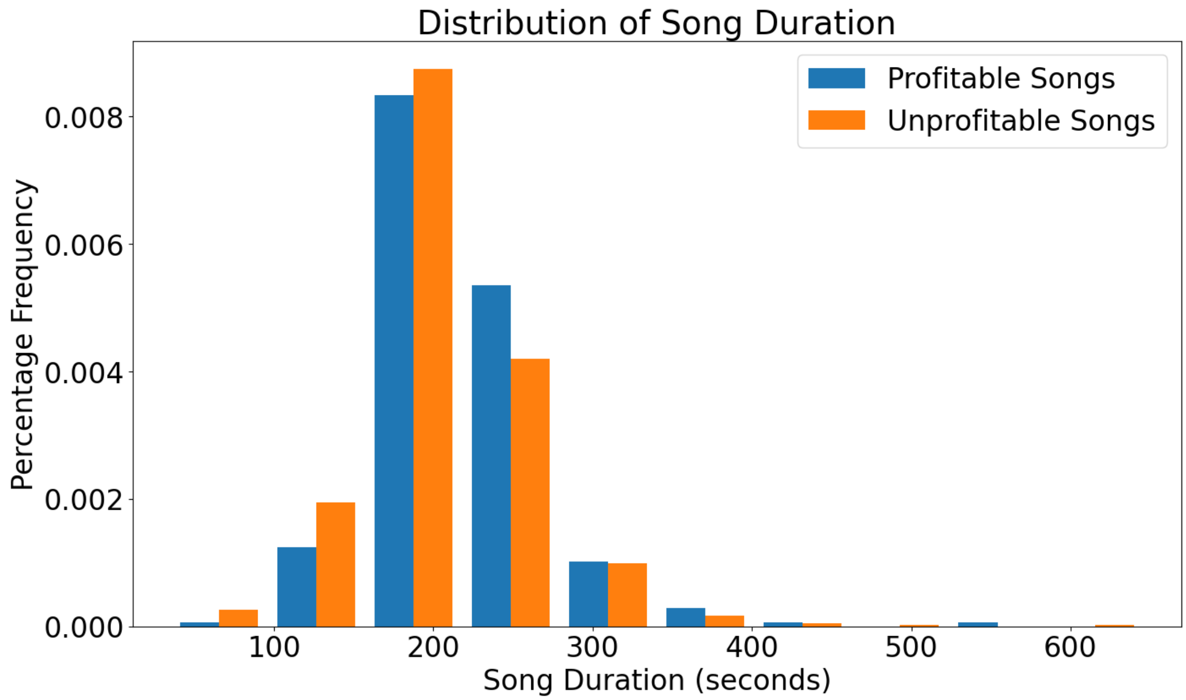

Figure 18: Counterfactual Duration - Popular Recommendations

Unprofitable songs have the same length as profitable songs under popular recommendations.

popular recommendations help some songs and hurt others. On average, however, songs are worse off when popular recommendations are used. Indeed, of the 1053 songs I observe that were released in the first three quarters of 2018, only 291 (27.6%) are profitable. For those songs that are profitable, their gross profit margin is 41.4%, compared to the average of 20% observed in the industry.

I now turn to some song characteristic results and welfare implications of my counterfactual analysis. Figure 18 reports the average duration of songs between profitable and unprofitable songs.

The average duration of songs of profitable songs is 217 seconds, and the average duration of unprofitable songs is 205 seconds. This difference is significant at the 1% level.

Table 14 reports the average values of other song characteristics for profitable and unprofitable songs, the difference in means, and the difference in distributions (as evaluated by a KS-Test):

The data shows more modest differences between profitable and unprofitable songs under popular recommendations. The most notable difference remains duration, with profitable songs being approximately 12 seconds longer on average (217.3 vs 205.5 seconds), and the Kolmogorov-Smirnov test ($p = 0.0091$) confirming significantly different

Feature	Profitable	Unprofitable	Diff	KS-Test P-Value
Duration (s)	217.318	205.454	-11.864***	0.0091
	(2.986)	(1.929)		
Tempo	123.786	124.777	0.991	0.0505
	(1.646)	(1.122)		
Energy	0.609	0.639	0.030***	0.0594
	(0.010)	(0.005)		
Danceability	0.712	0.689	-0.023**	0.0004
	(0.009)	(0.005)		
Valence	0.444	0.436	-0.008	0.3223
	(0.012)	(0.008)		
Acousticness	0.211	0.206	-0.005	0.4172
	(0.014)	(0.008)		
Instrumentalness	0.009	0.008	-0.000	0.2917
	(0.003)	(0.002)		
Liveness	0.178	0.181	0.002	0.3646
	(0.007)	(0.005)		
Speechiness	0.150	0.166	0.016*	0.4404
	(0.007)	(0.005)		
Loudness	-6.441	-6.370	0.070	0.9672
	(0.148)	(0.084)		
Mode	0.643	0.578	-0.064	0.3309
	(0.028)	(0.018)		

Table 14: Counterfactual Song Characteristics - Popular Recommendations

Unprofitable songs are shorter, less danceable, and more energetic.

distributions of song lengths. Danceability also shows a statistically significant difference both in means and distributions ($p = 0.0004$), though the economic significance is relatively small with profitable songs scoring 0.712 versus 0.689 for unprofitable ones. Energy levels show a significant mean difference of 0.030 higher for unprofitable songs, though the distributional difference is marginally significant ($p = 0.0594$).

Notably, many musical characteristics show no significant differences in either means or distributions between profitable and unprofitable songs. Features such as valence, acousticness, instrumentalness, liveness, and loudness all have high KS test p-values (ranging from 0.32 to 0.97), suggesting very similar distributions between the two groups. Even tempo, which differs by about 1 BPM, shows only marginally significant distributional differences ($p = 0.0505$). Overall, popular recommendations appear to have a more modest impact on song characteristics compared to random recommendations, with only duration and danceability showing significant differences between profitable and unprofitable

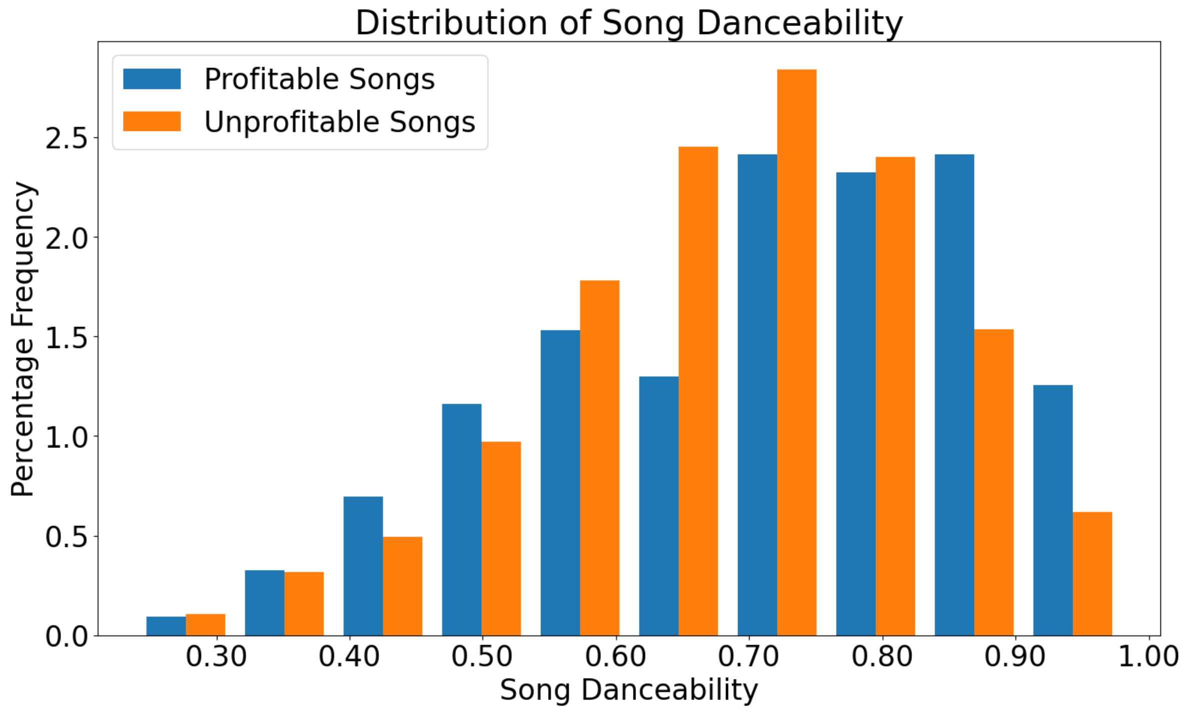

Figure 19: Counterfactual Danceability - Popular Recommendations

Unprofitable songs are less danceable than profitable songs under popular recommendations.

songs.

Figure 19 plots the distribution of song danceability for profitable and unprofitable songs:

The profitable songs are more danceable than unprofitable songs, and the entire distribution of danceability is right-shifted for profitable songs.

Finally, I turn to the welfare implications of my counterfactual analysis. I find that consumer surplus is 14.6% higher when targeted recommender systems are used, compared to popular recommendations. Restated, random recommender systems result in a 12.9% decrease in consumer surplus.

8 Conclusion

As recommender systems become increasingly integrated into the US economy, it is paramount that we understand their impact on both consumer demand and equilibrium supply decisions. This paper builds a structural model of the music streaming industry to estimate how these systems have influenced music production and shaped the sound of popular music. Because the music industry has historically been at the vanguard of technological adoption and disruption, understanding the impact of recommender systems on music releases can yield valuable insights into how these systems may affect other content-driven industries such as film, television, and digital commerce.

Using rich data on streaming sessions and song characteristics, I find that recommender systems have indeed changed the sound of music. The introduction of algorithmic recommendations has incentivized the release of shorter, more homogeneous songs optimized for the platform's objectives. However, these changes come with important welfare implications: while recommender systems have rendered songs more uniform in some dimensions, they have also enabled more songs to profitably enter the market and find their audience. My estimates suggest that this has increased consumer surplus by approximately 4% compared with a random recommendation system.

A key finding is the misalignment between recommender system preferences and consumer preferences. This partly stems from divergent incentives between platforms and rightsholders: while rightsholders want consumers to listen to at least 30 seconds of their songs to earn royalties, Spotify's recommender system is optimized for complete listens to reduce royalty payouts. In counterfactual analyses without algorithmic recommendations, I find that songs align more closely with raw consumer preferences, though fewer songs are profitable enough to be released.

My analysis of alternative recommendation schemes reveals important trade-offs. A popularity-based system that recommends songs based purely on aggregate listening statistics would generate a strong superstar effect: although it would make some songs highly profitable, it reduces overall consumer surplus by 13% compared with personalized recommendations. This suggests that while current recommender systems may reduce musical diversity in some dimensions, they help surface a wider variety of songs that appeal to different consumer tastes.

These findings have significant implications for platform regulation and competition policy. As regulators scrutinize digital platforms' use of algorithms, my results highlight both the efficiency gains from personalized recommendations and their potential to shape creative production. The framework developed here could inform antitrust analysis of other platforms on which algorithmic recommendations mediate between producers and consumers.

Several promising avenues exist for future research. First, extending this model to other content platforms, particularly short-form video services like TikTok, in which consumers directly choose to skip or watch videos and content is almost entirely algorithm-driven, could yield insights into how recommender systems shape creative production more broadly. Second, incorporating random coefficients would enrich the consumer demand structure and better capture preference heterogeneity. Third, endogenizing platform decisions with respect to pricing and recommendation algorithms would allow for deeper analysis of platform market power and the strategic manipulation of recommendations. Finally, examining how recommender systems affect market concentration and creator inequality remains an important question.

This research contributes to our understanding of how algorithmic systems shape markets and creative production in the digital economy. As these systems become more prevalent across industries, the methods and insights developed here can inform both managerial decisions and policy discussions regarding the regulation of digital platforms and their recommendation algorithms.

Appendix 1: Robustness Check of Demand Preferences over Time

Throughout this paper, I assume that consumer preferences are fixed over time. It is reasonable to claim, however, that these preferences can fluctuate over time, and that firms are responding to these fluctuations as well as the recommender system. To test this assumption, I conduct two robustness checks on consumer-preferences: a reduced-form difference-in-differences analysis, and a discrete choice model with time-varying coefficients.

For both of these analyses, I use my Spotify charts data, and examine the choice to listen as a function of song characteristics and time fixed effects. In my reduced-form specification, I interact song length with a time trend, to see the impact of these variables on the number of streams a song receives. In my discrete choice model, I assume consumers choose one song on the Spotify charts to listen to, and I estimate the probability they listen to a song as a function of song characteristics and time fixed effects.

8.1 Reduced Form Analysis

I estimate the following equation:

$$\begin{aligned} log(\text{Streams}_{jt}) = \alpha + \beta_1 \text{Duration}_j + \beta_2 \text{Time Trend}_t \\ + \delta(\text{Duration} \times \text{Time Trend})_{jt} + \gamma X_j + \eta_t + \epsilon_{jt} \end{aligned} \tag{14}$$

Here, Streams_{jt} is the number of streams song j receives on day t, Duration_j is the duration of song j, and Time Trend_t is the time trend for day t. Our coefficient of interest is δ, which captures the impact of song length on streams over time. I control for other song characteristic and week-of-year fixed effects.

Table 15 reports the results of this regression:

I find that the coefficient on the interaction is positive and significant at the 1% level, suggesting that consumer preferences are changing over time. Specifically, this result suggests that consumers are becoming more likely to listen to longer songs over time. This effect, however, is not economically meaningful. The coefficient on the interaction term is 0.00002, suggesting that a one-day change in the data, holding duration constant, increases streams by 0.002%. From the beginning to the end of the five-year sample period, this effect only amounts to an approximately 3% increase in streams.

This analysis, however, does not control for the growth in Spotify's user base, which

	Dependent variable: $log(Streams)$
Intercept	12.605***
	(0.008)
Duration	-0.027***
	(0.002)
Duration^2	0.0001
	(0.000)
Time Trend	0.00003***
	(0.000)
Duration × Time Trend	0.00002***
	(0.000)
Observations	364,081
R^2	0.019
Adjusted R^2	0.019
F Statistic	97.228*** (df=73; 364007)
Note:	*p<0.1; **p<0.05; ***p<0.01
	Other song characteristics and week fixed effects omitted for brevity. See Appendix

Table 15: Difference-in-Differences Results

Consumer preferences are changing over time, but not at an economically meaningful rate.

could also be driving this effect. A demand model with time-varying coefficients can better control for this effect.

8.2 Discrete Choice Model

I construct a discrete-choice model where consumers choose one song on Spotify to listen to. They can choose from among the top 50 songs on Spotify in a given week, with any songs outside the top 50 (positions 50-200) being an outside option. This captures choice on Spotify's Weekly Top 50 chart.

Consumers have the following utility function:

$$\begin{aligned} U_{ijti} = \alpha + \beta_1 \text{Duration}_j + \\ + \delta(\text{Duration} \times \text{Time Trend}_{jt}) + \gamma X_j + \eta_t + \epsilon_{ijt} \end{aligned} \tag{15}$$

Here, U_{ijt} is the utility consumer i receives from listening to song j on day t, and β_{i1} is the preference for song length. As before, δ captures the impact of song length on streams over time. My other control variable includes month fixed effects, to control for

	Dependent variable: $log(Market\ Share) - log(Outside\ Share)$
Duration	-0.487^{***}
	(0.059)
Duration^2	0.006
	(0.005)
Duration × Time Trend	0.003^{***}
	(0.0003)
Observations	10,350
Note:	*p<0.1; **p<0.05; ***p<0.01
Other song characteristics and month fixed effects omitted for brevity. See Appendix	

Table 16: Discrete Choice Model Results

Consumer preferences are changing over time, but not at an economically meaningful rate.

seasonality in music listening.

I estimate this model using PyBLP, instrumenting duration with characteristic of rival songs.

Table 16 reports the results of this regression:

Similar to the difference-and-difference analysis, I find that the coefficient on the interaction term positive, significant, but not economically meaningful. This coefficient has a less direct interpretation, as it is part of a discrete choice model, rather than a reduced-form regression.

In both cases, I find that consumer preferences for song duration are increasing over time, but not at an economically meaningful rate. Additionally, this movement is positive, rather than negative, suggesting that the trend towards shorter songs is not driven by consumer preferences, but rather by other factors. This result suggests that the model's assumption of fixed consumer preferences is reasonable, and that the model is capturing the impact of the recommender system on song releases.

Appendix 2: Nested Logit Specification

In many papers in industrial organization, the researcher specifies a nested logit model, with the outside option as its own nest. Such a choice structure would look like the following figure:

I estimate a conditional logit and nested logit model to compare the two. Following Reimers and Waldfogel (2023), I estimate the nested logit model in a bottom-up fashion,

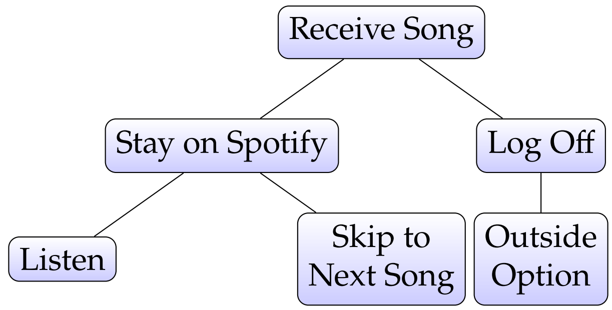

Figure 20: Nested Consumer Decision Tree

The consumer first chooses whether to stay on Spotify, then whether to skip or listen to a song.

estimating the inside options first, then the nest parameter.

Tables 17 and 18 reports consumer demand estimates for both the conditional logit and nested logit structure for consumers using the recommender system:

The coefficients in each model are identical. Additionally, the nested logit parameter, σ, is one, suggesting that the nested logit collapses into a conditional logit model.

Appendix 3: Oracular Recommender Counterfactual

The second counterfactual analysis I conduct is an oracular recommender system. I define an oracular recommender as one where the recommender is capable of giving the best possible song to each consumer, according to each consumer's preferences. Such a recommender system tends not to be feasible for several reasons: insufficient data, the cost of specifying such a granular model, and countervailing financial incentives. Bourreau and Gaudin (2022) and Reimers and Waldfogel (2023) both describe models in which platforms have incentives to bias recommender systems to maximize their own profit.

I implement this counterfactual in the following way:

1. Draw 10000 consumers and give them preferences from the demand estimates. I sample from the distribution of consumer characteristics in the data to construct values for those parameters.

2. Simulate a streaming session of 20 songs for each consumer, drawing from songs in the release window.

3. Compute the consumer surplus of this session, as well as the average song characteristics

	Dependent variable: Royalty Bearing Stream (RBS)			
	Direct Selection		Recommender System Selection	
	Estimate (Robust Std. Error)	Odds Ratio	Estimate (Robust Std. Error)	Odds Ratio
Age	0.00161	1.000	0.00118	1.000
	(0.00168)		(0.00168)	
Acousticness	0.0423***	0.689	0.0420***	0.688
	(0.00196)		(0.00197)	
Acousticness^2	-0.0378***	–	-0.0378***	–
	(0.000252)		(0.000253)	
Danceability	0.150***	3.287	0.150***	3.287
	(0.00362)		(0.00385)	
Danceability^2	0.0651***	–	0.0651***	–
	(0.00124)		(0.00124)	
Duration	-0.00254	1.000	-0.00159	1.000
	(0.00296)		(0.00297)	
Duration^2	0.00759***	–	0.00758***	–
	(0.000446)		(0.000447)	
Energy	-0.0648***	4.552	-0.0653***	4.545
	(0.00437)		(0.00444)	
Energy^2	0.161***	–	0.161***	–
	(0.00106)		(0.00109)	
Instrumentalness	-0.0368***	0.672	-0.0372***	0.671
	(0.00158)		(0.00167)	
Instrumentalness^2	-0.00628***	–	-0.00628***	–
	(0.000111)		(0.000111)	
Liveness	0.00530***	1.061	0.00535***	1.060
	(0.00177)		(0.00178)	
Liveness^2	0.000740***	–	0.000702***	–
	(0.000168)		(0.000168)	
Loudness	0.0488***	1.010	0.0483***	1.010
	(0.00355)		(0.00358)	
Loudness^2	-0.0136***	–	-0.0136***	–
	(0.000361)		(0.000366)	
Mode	0.0484***	1.050	0.0484***	1.050
	(0.00313)		(0.00315)	
μ	–	–	1***	–
	–		(0.00927)	
Speechiness	-0.00170	0.629	-0.00158	0.629
	(0.00170)		(0.00172)	
Speechiness^2	-0.00829***	–	-0.00830***	–
	(0.000149)		(0.000156)	
Tempo	0.0184***	1.000	0.0192***	1.000
	(0.00289)		(0.00296)	
Tempo^2	0.00976***	–	0.00972***	–
	(0.00113)		(0.00113)	
Time Signature	-0.0597***	0.942	-0.0545***	0.947
	(0.0109)		(0.0109)	
Valence	0.0155***	0.995	0.0158***	0.996
	(0.00234)		(0.00235)	
Valence^2	-0.00517***	–	-0.00517***	–
	(0.00059)		(0.000591)	
Model Statistics				
Observations	31,238,428		31,238,428	
$\bar{\rho}^2$	0.284		0.284	

Notes:
*** p<0.01, ** p<0.05, * p<0.1

Table 17: Logit and Nested Logit Demand Estimates - Song Characteristics

	Dependent variable: Royalty Bearing Stream (RBS)			
	Direct Selection		Recommender System Selection	
	Estimate (Robust Std. Error)	Odds Ratio	Estimate (Robust Std. Error)	Odds Ratio
Time of Day				
Morning	0.150***	1.162	0.151***	1.163
	(0.00415)		(0.00423)	
Afternoon	0.121***	1.129	0.122***	1.130
	(0.00371)		(0.00376)	
Night	0.125***	1.133	0.124***	1.132
	(0.00611)		(0.00613)	
Day of Week				
Tuesday	0.0289***	1.029	0.0296***	1.030
	(0.00538)		(0.00538)	
Wednesday	0.0231***	1.023	0.0246***	1.025
	(0.00552)		(0.00553)	
Thursday	0.0207***	1.021	0.0225***	1.023
	(0.00551)		(0.00552)	
Friday	-0.0520***	0.949	-0.0508***	0.950
	(0.00529)		(0.00532)	
Saturday	-0.0220***	0.978	-0.0211***	0.979
	(0.00568)		(0.00568)	
Sunday	-0.0117**	0.988	-0.0107*	0.989
	(0.00566)		(0.00566)	
User Characteristics				
Premium	0.116***	1.123	0.115***	1.122
	(0.00465)		(0.00467)	
Model Statistics				
Observations	31,238,428		31,238,428	
$\bar{\rho}^2$	0.284		0.292	

Notes:
*** p<0.01, ** p<0.05, * p<0.1

Table 18: Logit and Nested Logit Demand Estimates - Contextual Characteristics

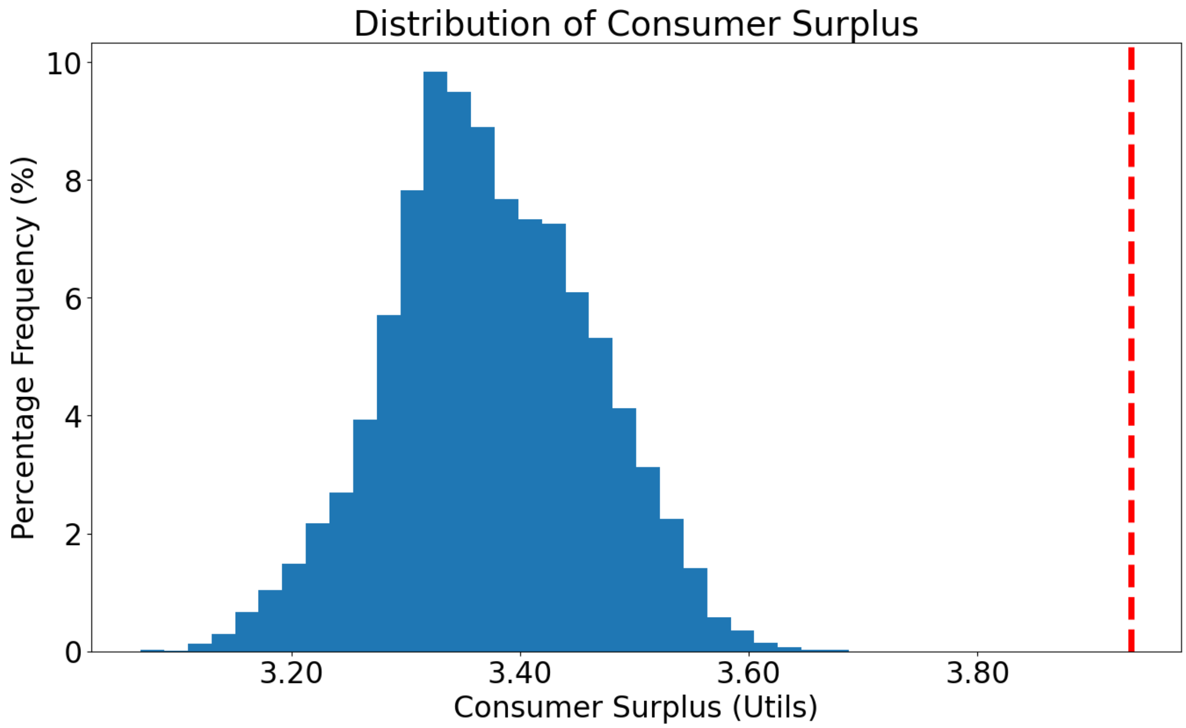

Figure 21: Counterfactual Consumer Surplus - Oracular Recommendations

The oracular recommender increases consumer surplus by 5.3% compared to the simulated streaming sessions.

4. Compute the consumer surplus of the 20 highest-utility songs, as well as the average song characteristics of those songs.

5. Compare results between the two sets of songs.

First, I compare consumer surplus generated by these streaming sessions. Figure 21 reports the results of this comparison:

Each blue bar represent streaming sessions, and the red line represents the utility-maximizing streaming session. The oracular recommender increases consumer surplus by 16.6% compared to the simulated streaming sessions. This difference is statistically significant at the 1% level.

Table 19 reports the average values of song characteristics for the simulated streaming sessions and the utility-maximizing streaming sessions, and the difference in means:

Songs in the optimal streaming session is more acoustic, more danceable, and less energetic than the simulated sessions. This suggests that the oracular recommender system is more likely to recommend songs slower songs than a random recommender.

Characteristic	Mean (Simulated)	Mean (Optimal)	Difference
Acousticness	0.20	0.40	-0.19***
			(-3.83)
Age	0.05	0.15	-0.10
			(-1.38)
Danceability	0.70	0.76	-0.06*
			(-1.86)
Duration	207.82	184.65	23.17**
			(2.00)
Energy	0.63	0.29	0.34***
			(9.67)
Instrumentalness	0.01	0.00	0.01
			(0.60)
Liveness	0.18	0.19	-0.01
			(-0.45)
Loudness	-6.38	-9.68	3.31***
			(6.37)
Speechiness	0.16	0.16	0.00
			(0.01)
Tempo	124.22	112.95	11.27*
			(1.65)
Valence	0.44	0.42	0.02
			(0.40)
Consumer Surplus	3.37	3.93	-0.56***
			(-6.35)
Note:	T-statistics in parentheses; *p<0.1; **p<0.05; ***p<0.01		

Table 19: Counterfactual Song Characteristics and Consumer Surplus - Oracular Recommendations

The oracular recommender system surfaces more acoustic, more danceable, and less energetic than the simulated streaming sessions.

Supplemental Tables

Characteristic	Description
Acousticness	A confidence measure from 0.0 to 1.0 of whether the track is acoustic. 1.0 represents high confidence the track is acoustic.
Danceability	Danceability describes how suitable a track is for dancing based on a combination of musical elements including tempo, rhythm stability, beat strength, and overall regularity. A value of 0.0 is least danceable and 1.0 is most danceable.
Energy	Energy is a measure from 0.0 to 1.0 and represents a perceptual measure of intensity and activity. Typically, energetic tracks feel fast, loud, and noisy. For example, death metal has high energy, while a Bach prelude scores low on the scale. Perceptual features contributing to this attribute include dynamic range, perceived loudness, timbre, onset rate, and general entropy.
Instrumentalness	Predicts whether a track contains no vocals. "Ooh" and "aah" sounds are treated as instrumental in this context. Rap or spoken word tracks are clearly "vocal". The closer the instrumentalness value is to 1.0, the greater likelihood the track contains no vocal content. Values above 0.5 are intended to represent instrumental tracks, but confidence is higher as the value approaches 1.0.
Liveness	Detects the presence of an audience in the recording. Higher liveness values represent an increased probability that the track was performed live. A value above 0.8 provides strong likelihood that the track is live.
Speechiness	Speechiness detects the presence of spoken words in a track. The more exclusively speech-like the recording (e.g. talk show, audiobook, poetry), the closer to 1.0 the attribute value. Values above 0.66 describe tracks that are probably made entirely of spoken words. Values between 0.33 and 0.66 describe tracks that may contain both music and speech, either in sections or layered, including such cases as rap music. Values below 0.33 most likely represent music and other non-speech-like tracks.
Valence	A measure from 0.0 to 1.0 describing the musical positiveness conveyed by a track. Tracks with high valence sound more positive (e.g. happy, cheerful, euphoric), while tracks with low valence sound more negative (e.g. sad, depressed, angry).

Table 20: Descriptions of Song Characteristics

	Estimate (Robust Std. Error)	Odds Ratio
Age	-0.0161***	0.998
	(0.00047)	
Acousticness	0.0581***	0.680
	(0.000601)	
Acousticness^2	-0.043***	–
	(0.000101)	
Danceability	-0.00696***	0.966
	(0.00108)	
Danceability^2	-0.00107**	–
	(0.000442)	
Duration	0.00991***	1.000
	(0.000821)	
Duration^2	-0.000538***	–
	(0.0000838)	
Energy	-0.0104***	1.856
	(0.00126)	
Energy^2	0.0608***	–
	(0.000371)	
Instrumentalness	-0.00516***	0.722
	(0.000495)	
Instrumentalness^2	-0.00960***	–
	(0.0000438)	
Liveness	0.0107***	0.844
	(0.000512)	
Liveness^2	-0.00635***	–
	(0.0000626)	
Loudness	0.00734***	1.001
	(0.000988)	
Loudness^2	-0.00802***	–
	(0.0000992)	
Mode	0.0332***	1.034
	(0.000965)	
Speechiness	-0.00594***	0.544
	(0.00049)	
Speechiness^2	-0.0104***	–
	(0.0000562)	
Tempo	-0.00264***	1.000
	(0.000891)	
Tempo^2	-0.00168***	–
	(0.000441)	
Time Signature	0.0334***	1.034
	(0.00305)	
Valence	0.0272***	0.596
	(0.000714)	
Valence^2	-0.0566***	–
	(0.000239)	

Notes:
*** p<0.01, ** p<0.05, * p<0.1

Table 21: Full Sample Consumer Demand Estimates - Song Characteristics

	Estimate (Robust Std. Error)	Odds Ratio
Time of Day		
Morning	0.139***	1.149
	(0.00129)	
Afternoon	0.110***	1.116
	(0.00116)	
Night	0.0526***	1.054
	(0.00182)	
Day of Week		
Tuesday	0.00859***	1.009
	(0.00170)	
Wednesday	0.0118***	1.012
	(0.00174)	
Thursday	0.0039**	1.004
	(0.00173)	
Friday	-0.00988***	0.990
	(0.00169)	
Saturday	-0.0163***	0.984
	(0.00176)	
Sunday	-0.0159***	0.984
	(0.00175)	
User Characteristics		
Premium	-0.0708***	0.932
	(0.00134)	

Notes:
*** p<0.01, ** p<0.05, * p<0.1

Table 22: Full Sample Consumer Demand Estimates - Contextual Characteristics

	Acousticness	Age	Danceability	Duration	Energy	Instrumentalness	Liveness	Loudness	Mode	Speechiness	Tempo	Time Signature	Valence
Constant	0.447**	-0.325	-0.320*	0.028	-0.320**	-0.131***	-0.104	-0.015	0.061	-0.044	-0.122	0.158***	-0.190*
	(0.144)	(0.265)	(0.127)	(0.075)	(0.111)	(0.040)	(0.064)	(0.122)	(0.040)	(0.095)	(0.069)	(0.014)	(0.086)
Drift	0.000**	0.000	-0.000***	-0.000***	-0.000	0.000***	0.000***	-0.000	0.000*	-0.000***	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Acousticness$_{t-1}$	**0.927*****	0.037	-0.051*	-0.080***	-0.028	-0.002	0.030*	-0.028	0.020**	0.017	0.025*	-0.009***	-0.052***
	(0.027)	(0.049)	(0.024)	(0.014)	(0.021)	(0.007)	(0.012)	(0.023)	(0.007)	(0.018)	(0.013)	(0.003)	(0.016)
Age$_{t-1}$	0.054***	**0.832*****	-0.025	-0.021**	-0.049***	0.008	-0.007	-0.018	0.000	-0.008	-0.016*	0.002	-0.061***
	(0.015)	**(0.028)**	(0.014)	(0.008)	(0.012)	(0.004)	(0.007)	(0.013)	(0.004)	(0.010)	(0.007)	(0.001)	(0.009)
Danceability$_{t-1}$	0.116***	0.008	**0.784*****	-0.061***	-0.091***	-0.002	0.002	-0.052*	0.029***	-0.035*	-0.035**	0.005	-0.083***
	(0.026)	(0.048)	**(0.023)**	(0.014)	(0.020)	(0.007)	(0.012)	(0.022)	(0.007)	(0.017)	(0.013)	(0.002)	(0.016)
Duration$_{t-1}$	0.029	0.022	-0.119***	**0.907*****	-0.039*	0.010	0.014	-0.042*	0.024***	-0.023	-0.008	-0.004	-0.046***
	(0.022)	(0.040)	(0.019)	**(0.011)**	(0.017)	(0.006)	(0.010)	(0.019)	(0.006)	(0.014)	(0.011)	(0.002)	(0.013)
Energy$_{t-1}$	0.078*	-0.040	-0.107***	-0.028	**0.858*****	-0.015	0.001	-0.016	0.036***	-0.053*	0.004	0.006	-0.122***
	(0.036)	(0.066)	(0.032)	(0.019)	**(0.028)**	(0.010)	(0.016)	(0.030)	(0.010)	(0.024)	(0.017)	(0.003)	(0.021)
Instrumentalness$_{t-1}$	0.019	-0.164	0.055	0.008	-0.046	**0.815*****	-0.005	0.007	0.002	0.059	-0.087***	0.012*	-0.072*
	(0.053)	(0.097)	(0.046)	(0.027)	(0.041)	**(0.015)**	(0.023)	(0.044)	(0.015)	(0.035)	(0.025)	(0.005)	(0.032)
Liveness$_{t-1}$	0.064*	0.120*	-0.017	0.001	-0.061**	0.006	**0.870*****	-0.071**	0.010	0.031	0.022	0.001	0.042*
	(0.031)	(0.056)	(0.027)	(0.016)	(0.024)	(0.009)	**(0.014)**	(0.026)	(0.009)	(0.020)	(0.015)	(0.003)	(0.018)
Loudness$_{t-1}$	-0.027	-0.059	0.086**	-0.038*	-0.013	0.004	-0.006	**0.900*****	-0.022*	0.091***	0.002	-0.010**	0.013
	(0.037)	(0.069)	(0.033)	(0.019)	(0.029)	(0.010)	(0.017)	**(0.032)**	(0.010)	(0.025)	(0.018)	(0.004)	(0.022)
Mode$_{t-1}$	-0.064	0.065	0.061	0.056**	0.103***	0.013	0.007	0.059	**0.918*****	-0.017	-0.005	0.005	0.075**
	(0.041)	(0.075)	(0.036)	(0.021)	(0.032)	(0.011)	(0.018)	(0.035)	**(0.011)**	(0.027)	(0.020)	(0.004)	(0.024)
Speechiness$_{t-1}$	-0.001	-0.094*	0.023	-0.004	0.003	0.014	0.015	0.030	-0.011	**0.887*****	-0.019	-0.003	-0.021
	(0.025)	(0.046)	(0.022)	(0.013)	(0.019)	(0.007)	(0.011)	(0.021)	(0.007)	**(0.017)**	(0.012)	(0.002)	(0.015)
Tempo$_{t-1}$	0.104***	0.093*	-0.090***	0.027*	-0.053**	-0.002	-0.006	-0.051*	0.007	-0.064***	**0.860*****	0.001	-0.019
	(0.025)	(0.045)	(0.022)	(0.013)	(0.019)	(0.007)	(0.011)	(0.021)	(0.007)	(0.016)	**(0.012)**	(0.002)	(0.015)
Time Signature$_{t-1}$	-0.377**	0.119	0.309*	-0.071	0.197	0.092*	0.068	-0.028	-0.010	0.065	0.110	**0.841*****	0.075
	(0.143)	(0.262)	(0.126)	(0.074)	(0.110)	(0.040)	(0.063)	(0.120)	(0.040)	(0.094)	(0.068)	**(0.013)**	(0.085)
Valence$_{t-1}$	0.018	0.028	-0.038*	-0.005	-0.036*	-0.009	0.007	-0.026	0.000	-0.010	-0.033***	-0.007***	**0.924*****
	(0.019)	(0.036)	(0.017)	(0.010)	(0.015)	(0.005)	(0.009)	(0.016)	(0.005)	(0.013)	(0.009)	(0.002)	**(0.012)**
Observations	1820	1820	1820	1820	1820	1820	1820	1820	1820	1820	1820	1820	1820
AIC	-100	-100	-100	-100	-100	-100	-100	-100	-100	-100	-100	-100	-100
BIC	-99	-99	-99	-99	-99	-99	-99	-99	-99	-99	-99	-99	-99

Note: Standard errors in parentheses. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. Own-characteristic lag terms are in bold.

Table 23: Vector Autoregression (VAR) Model Results

The VAR model suggests strong, stationary processes for each song characteristic with respect to its own lag. Most of the cross-characteristic lags are statistically insignificant.

Bibliography

References

Aguiar, Luis, and Joel Waldfogel. 2018. "Quality Predictability and the Welfare Benefits from New Products: Evidence from the Digitization of Recorded Music." Publisher: The University of Chicago Press, *Journal of Political Economy* 126, no. 2 (April 1, 2018): 492–524. ISSN: 0022-3808, accessed November 28, 2021. https://doi.org/10.1086/696229. https://www.journals.uchicago.edu/doi/full/10.1086/696229.

———. 2021. "Platforms, Power, and Promotion: Evidence from Spotify Playlists*." *The Journal of Industrial Economics* 69 (3): 653–691. ISSN: 1467-6451, accessed December 20, 2021. https://doi.org/10.1111/joie.12263. https://onlinelibrary.wiley.com/doi/abs/10.1111/joie.12263.

Aguiar, Luis, Joel Waldfogel, and Sarah Waldfogel. 2021. "Playlisting favorites: Measuring platform bias in the music industry." *International Journal of Industrial Organization* 78 (September 1, 2021): 102765. ISSN: 0167-7187, accessed November 29, 2021. https://doi.org/10.1016/j.ijindorg.2021.102765. https://www.sciencedirect.com/science/article/pii/S0167718721000588.

Anderson, Simon P., Andre de Palma, and Jacques-Francois Thisse. 1992. *Discrete Choice Theory of Product Differentiation.* Cambridge, MA, USA: MIT Press, October 16, 1992. ISBN: 978-0-262-01128-0.

Aridor, Guy, Duarte Goncalves, Daniel Kluver, Ruoyan Kong, and Joseph Konstan. 2023. "The Economics of Recommender Systems: Evidence from a Field Experiment on MovieLens." In *Proceedings of the 24th ACM Conference on Economics and Computation,* 117. EC '23. New York, NY, USA: Association for Computing Machinery, July 7, 2023. ISBN: 9798400701047, accessed January 14, 2024. https://doi.org/10.1145/3580507.3597677. https://doi.org/10.1145/3580507.3597677.

Aridor, Guy, and Duarte Gonçalves. 2022. "Recommenders' originals: The welfare effects of the dual role of platforms as producers and recommender systems." *International Journal of Industrial Organization* 83 (July 1, 2022): 102845. ISSN: 0167-7187, accessed September 5, 2022. https://doi.org/10.1016/j.ijindorg.2022.102845. https://www.sciencedirect.com/science/article/pii/S0167718722000212.

Benner, Mary J., and Joel Waldfogel. 2016. "The Song Remains the Same? Technological Change and Positioning in the Recorded Music Industry." *Strategy Science* 1, no. 3 (September): 129–147. ISSN: 2333-2050, 2333-2077, accessed November 28, 2021. https://doi.org/10.1287/stsc.2016.0012. http://pubsonline.informs.org/doi/10.1287/stsc.2016.0012.

Bourreau, Marc, and Germain Gaudin. 2022. "Streaming platform and strategic recommendation bias." *Journal of Economics & Management Strategy* 31 (1): 25–47. ISSN: 1530-9134, accessed February 6, 2022. https://doi.org/10.1111/jems.12452. https://onlinelibrary.wiley.com/doi/abs/10.1111/jems.12452.

Brost, Brian, Rishabh Mehrotra, and Tristan Jehan. 2018. "The Music Streaming Sessions Dataset" (December 31, 2018). Accessed March 26, 2022. https://doi.org/10.48550/arXiv.1901.09851. https://arxiv.org/abs/1901.09851v2.

Chace, Zoe. 2011. "How Much Does It Cost To Make A Hit Song?" *NPR* (June 30, 2011). Accessed May 2, 2024. https://www.npr.org/sections/money/2011/07/05/137530847/how-much-does-it-cost-to-make-a-hit-song.

Eriksson, Maria, Rasmus Fleischer, Anna Johansson, Pelle Snickars, and Patrick Vonderau. 2019. *Spotify Teardown: Inside the Black Box of Streaming Music.* Cambridge, MA, USA: MIT Press, February 19, 2019. ISBN: 978-0-262-03890-4.

Goeree, Michelle Sovinsky. 2008. "Limited Information and Advertising in the U.S. Personal Computer Industry." _Eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA4158, *Econometrica* 76 (5): 1017–1074. ISSN: 1468-0262, accessed November 27, 2021. https://doi.org/10.3982/ECTA4158. https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA4158.

Lancaster, Kelvin J. 1966. "A New Approach to Consumer Theory." Publisher: University of Chicago Press, *Journal of Political Economy* 74 (2): 132–157. ISSN: 0022-3808, accessed April 12, 2022. https://www.jstor.org/stable/1828835.

Lee, Robin S. 2013. "Vertical Integration and Exclusivity in Platform and Two-Sided Markets." *American Economic Review* 103, no. 7 (December): 2960–3000. ISSN: 0002-8282, accessed April 13, 2023. https://doi.org/10.1257/aer.103.7.2960. https://www.aeaweb.org/articles?id=10.1257/aer.103.7.2960.

McInerney, James, Benjamin Lacker, Samantha Hansen, Karl Higley, Hugues Bouchard, Alois Gruson, and Rishabh Mehrotra. 2018. "Explore, exploit, and explain: personalizing explainable recommendations with bandits." In *Proceedings of the 12th ACM Conference on Recommender Systems,* 31–39. RecSys '18: Twelfth ACM Conference on Recommender Systems. Vancouver British Columbia Canada: ACM, September 27, 2018. ISBN: 978-1-4503-5901-6, accessed January 10, 2022. https://doi.org/10.1145/3240323.3240354. https://dl.acm.org/doi/10.1145/3240323.3240354.

Melchiorre, Alessandro B., Navid Rekabsaz, Emilia Parada-Cabaleiro, Stefan Brandl, Oleg Lesota, and Markus Schedl. 2021. "Investigating gender fairness of recommendation algorithms in the music domain." *Information Processing & Management* 58 (5): 102666. ISSN: 0306-4573. https://doi.org/https://doi.org/10.1016/j.ipm.2021.102666. https://www.sciencedirect.com/science/article/pii/S0306457321001540.

Reimers, Imke, and Joel Waldfogel. 2023. *A Framework for Detection, Measurement, and Welfare Analysis of Platform Bias,* 31766, October. Accessed October 9, 2023. https://doi.org/10.3386/w31766. https://www.nber.org/papers/w31766.

Singer, Howie, and Bill Rosenblatt. 2023. "Streaming: Islands in the Stream—Dolly Parton and Kenny Rogers." In *Key Changes: The Ten Times Technology Transformed the Music Industry,* edited by Howie Singer and Bill Rosenblatt, 0. Oxford University Press, October 22, 2023. ISBN: 978-0-19-765689-1, accessed January 12, 2024. https://doi.org/10.1093/oso/9780197656891.003.0009. https://doi.org/10.1093/oso/9780197656891.003.0009.

Singleton, Micah. 2015. "This was Sony Music's contract with Spotify." The Verge, May 19, 2015. Accessed April 15, 2024. https://www.theverge.com/2015/5/19/8621581/sony-music-spotify-contract.

Train, Kenneth E. 2009. *Discrete Choice Methods with Simulation.* 2nd ed. Cambridge: Cambridge University Press. ISBN: 978-0-521-76655-5, accessed April 10, 2024. https://doi.org/10.1017/CBO9780511805271. https://www.cambridge.org/core/books/discrete-choice-methods-with-simulation/49CABD00F3DDDA088A8FBFAAAD7E9546.

Weintraub, Gabriel Y., Lanier Benkard, and Benjamin Van Roy. 2005. "Oblivious Equilibrium: A Mean Field Approximation for Large-Scale Dynamic Games." In *Advances in Neural Information Processing Systems,* vol. 18. MIT Press. Accessed March 11, 2024. https://proceedings.neurips.cc/paper/2005/hash/fb3f76858cb38e5b7fd113e0bc1c0721-Abstract.html.