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

Max E. Schnidman†

September 26, 2024, latest version available here

Abstract

I examine how recommender systems have influenced the music industry and shaped music production over the last decade. Using a structural model of Spotify's recorded music industry, I analyze consumer behavior, platform recommendations, and rightsholder release decisions. Results indicate that streaming services and recommender systems correlate with a 40-second decrease in average song length on Billboard's Hot 100. The model estimates a fixed cost of $85,000 for songs entering Spotify's Top 200, with an 11% price-cost margin. Counterfactual analysis shows that without recommender systems, songs would be 10 seconds longer on average, but consumer welfare would be 9% lower. While reducing music variety, these systems have increased song quantity and overall 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, 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. All errors are my own.

†University of Virginia, mes7jw@virginia.edu

1 Introduction

Recommender systems, 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, where consumers can access a vast catalog of music for a fixed monthly fee, have become the primary way that consumers access music, with streaming accounting for 84 percent of the recorded music industry's $16bn revenue in 2023.[2] Notably, these platforms use recommender systems to generate algorithmic playlists to surface music for users, These playlists 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 over the last decade.

Recommender systems are a form of advertising for content on digital platforms, but they are unique in that the producer does not actually design or purchase the advertisement.[3] 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.[4] These systems come with a number of economic tradeoffs. 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). Platforms can use these algorithms to steer consumers towards profit-maximizing products, rather than products that consumers actually prefer (Reimers and Waldfogel 2023). Additionally, these systems may have inherent biases, providing 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, where

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

producers respond to the 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 form 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 listens to a particular song, based on the song's characteristics and the consumer's preferences, and delivers the song to the consumer. Rightsholders choose whether to release songs to Spotify given the demand for the song, which is the joint 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, which is given by the expected revenue of the worst-performing (or marginal) song.

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 160mn consumer-level streaming sessions from July to September 2018, including 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 webpage reporting the daily top 200 songs on Spotify for every country in which they operate. It also includes stream counts, and the song ID. 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. 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. Additionally, music industry executives have 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, and that producers respond to this gap by targeting the recommenders' objectives jointly with consumer preferences. For example, while consumer 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 music on Spotify. My estimate for a song that enters Spotify's Top 200 is $85,000, and my estimate for the average song is $41. In both cases, however, the price-cost margin is small, at approximately 11% for a song that enters the Top 200, and approximately 9% for the average song.

My counterfactual analysis focuses on changing Spotify's recommender system to see how it has affected song characteristics. Specifically, I impose random recommendations, as a proxy for no recommendations. I find that in the absence of recommender systems, songs are on average 10 seconds longer, more heterogeneous, and less profitable. As a result, fewer songs are released, and consumer welfare is 9% lower than in the status quo. 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.

The results from this model suggest that digital platforms can use their recommender systems strategically to affect both the demand for and supply of products on their platform. It also shows that consumers care more about the quantity of available products, even if they are more homogeneous, than the variety of products available, and that recommender systems can help improve consumer welfare. Additionally, it suggests that firms need to consider both the consumer and the platform when designing their products, and that they should be aware of the strategic implications of their decisions. Moreover, this research also has implications for antitrust authorities, who should consider the effects of these systems on competition and consumer welfare when evaluating mergers and acquisitions in the digital space.

This 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 I use in this paper 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, modifying the recommender system to observe how equilibrium song releases change. 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. Aguiar, Waldfogel, and Waldfogel 2021 uses 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, suggesting that the curators are looking to maximize ex-post streams. 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 ex-post performance would suggest. Benner and Waldfogel 2016 use a difference-in-difference design to estimate how 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 taking these insights into Spotify playlists and digitization and embedding them in a structural model of the industry. It also builds on Aguiar and Waldfogel 2018, which developed 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 different scenarios: perfect quality foresight, no quality foresight, and imperfect quality foresight. 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 find that tripling the choice set under imperfect foresight adds nearly 20 times as much consumer surplus as tripling the choice set under perfect foresight. I extend this model into the music streaming industry, by modifying the choice structure to reflect the streaming industry, incorporating a recommender system into the model, and introducing forward-looking rightsholders. I apply their entry condition to estimate the fixed cost of entry on Spotify.

Second, I contribute to a growing literature on recommender systems in economics. Many recent papers have focused on the effect of recommender systems on pricing. Calvano et al. 2020 embeds learning algorithms in a repeated Bertrand oligopoly setting, and they find that these algorithms result in supracompetitive prices. This outcome is a result of collusive strategies that the algorithms adapt without explicitly communicating with each other. I extend this paper's insight into equilibrium effects of algorithms into product characteristics. Other papers have studied the effects of recommender systems through theoretical models or through reduced-form analysis. Bourreau and Gaudin 2022 uses a Hotelling model of music listening with a recommender system and a digital platform hosting both songs. They find that the platform uses the recommender system to drive consumers to cheaper songs, even if they are further away from the consumer's ideal song. Aridor and Gonçalves 2022 similarly embeds 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 towards its own products, and that this can reduce consumer welfare through 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, focusing on how these systems affect producer product decisions. Melchiorre et al. 2021 introduces 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 the recommendations towards certain gender groups. Aridor et al. 2023 conducts 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 towards certain profit-maximizing products, at the expense of consumer welfare. Lee 2013 constructs an empirical model of the video games industry, focusing on the role of exclusive games on console platforms. He finds that in the absence of exclusivity agreements, console sales and consumer welfare would both be higher, but that only the incumbent console manufacturer would benefit from the removal of such agreements. I extend his model of games production to the music industry, and I 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 establishes a welfare frontier for platforms, which is a weighted sum of consumer and producer surplus. 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, finding that both platforms are off the frontier. Aguiar and Waldfogel 2021 estimate the effect of including a song on a Spotify playlist using a 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 into my model of the music industry.

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.

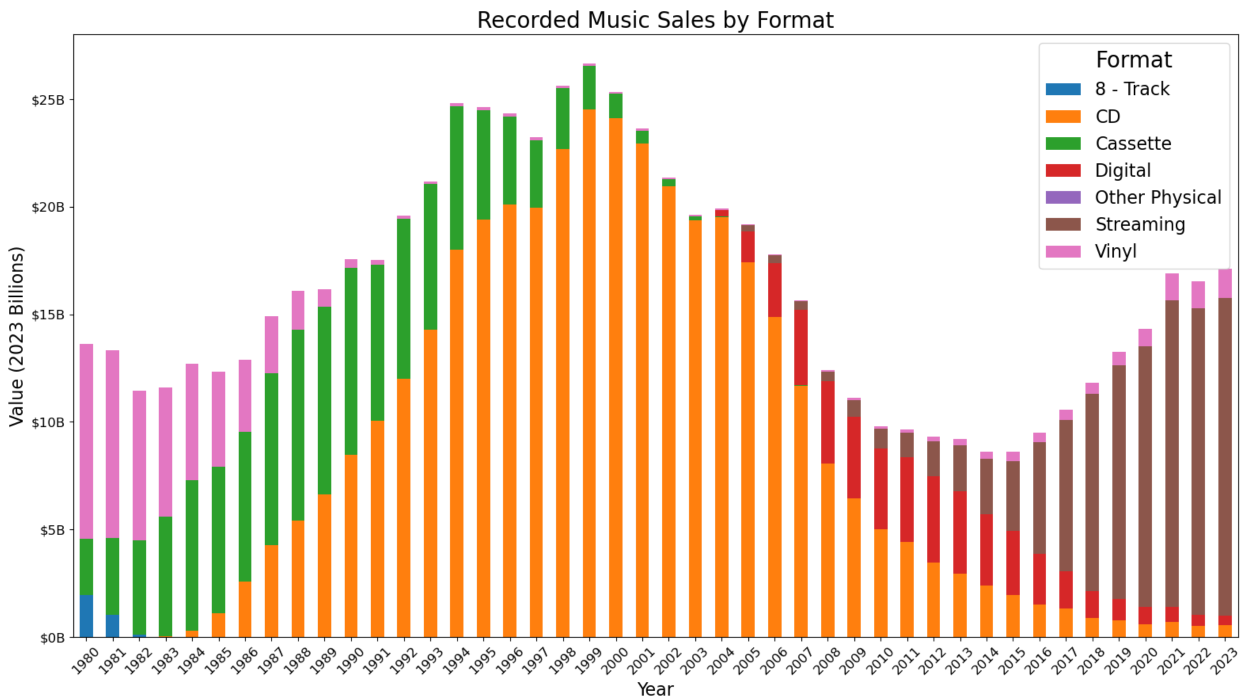

Figure 1: Real Revenue of the Recorded Music Industry, 1990-2023

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. The creation of these services was concurrent with a marked decline in revenues for the industry, which industry participants attributed to these illegal file-sharing services, and legal challenges brought by the Recording Industry Association of

America resulted in their closure in the 2000s.

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.[5] Additionally, Apple negotiated a revenue-sharing deal with labels, giving them 30 percent 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.[6] Such services did not see widespread acceptance until the early 2010s, when Spotify launched in the U.S. Spotify combined an expansive 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.[7] 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.[8]

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,

5. https://www.engadget.com/2013-04-29-the-itunes-influence-part-one.html
6. https://www.thurrott.com/music-videos/groove-music/6033/microsoft-is-finally-retiring-zune-zune-music-pass
7. https://www.theverge.com/2012/1/6/2688250/spotify-free-account-restriction-10-hours-per-month
8. 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

these services have become the primary way that consumers access music, with streaming accounting for 84 percent 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 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 16).

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.[9] 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:

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

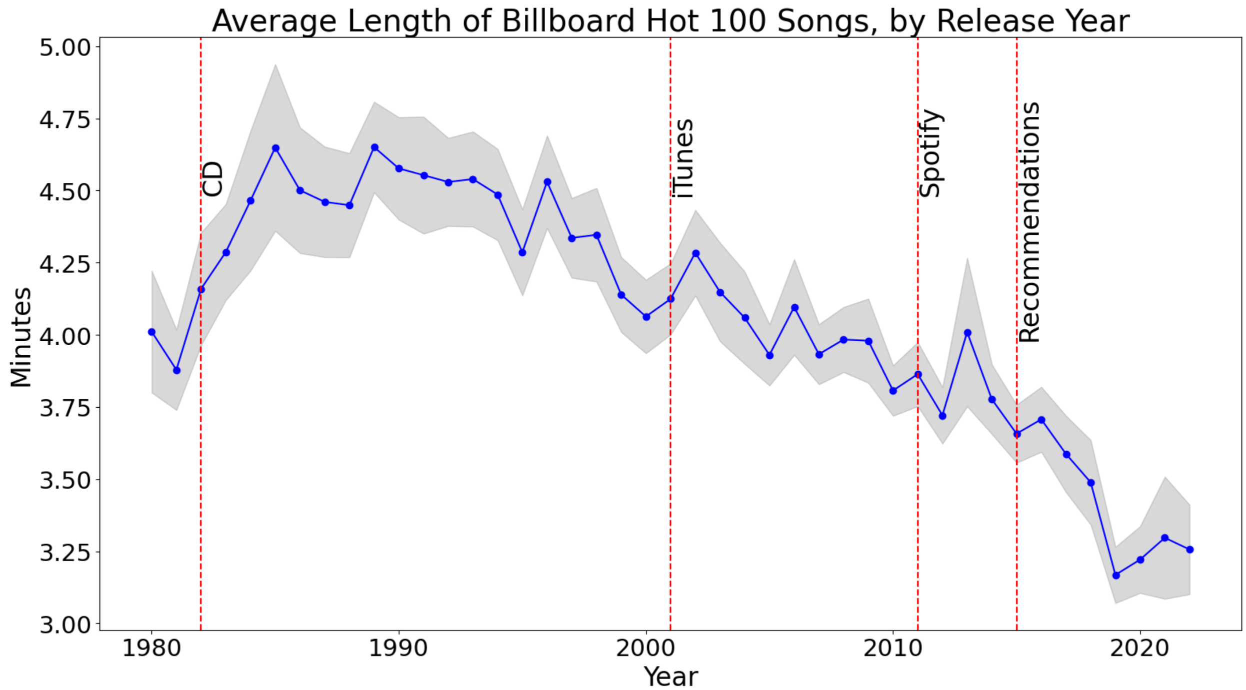

Figure 2: Average Song Duration on Billboard's Hot 100, 1990-2023

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

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 percent 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

	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

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.[10]

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.

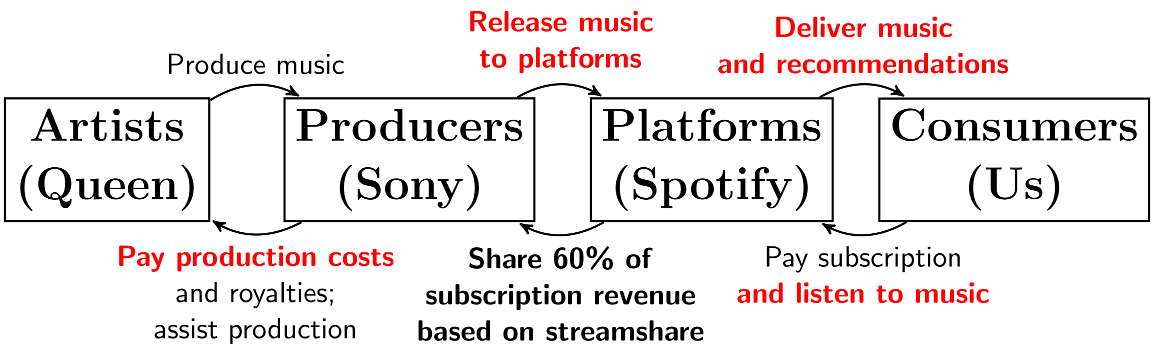

Figure 3: Vertical Structure in the Music Industry

Beginning on the left, artists create music, either by themselves or in contract with rightsholders. An artist on contract with a rightsholder (e.g., a record label) typically receives and 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.[11] The market for artists is highly diffuse, with tens of thousands of artists working on music each day, competing not just with 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.[12]

Rightsholders, such as Sony, Warner, and Universal (the Big Three record labels), are

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

11. 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

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

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 sublabels) 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 percent of the market. Other independent labels comprise the remaining 23 percent of the market. Figure 4 shows the market share of rightsholders (and streaming services).

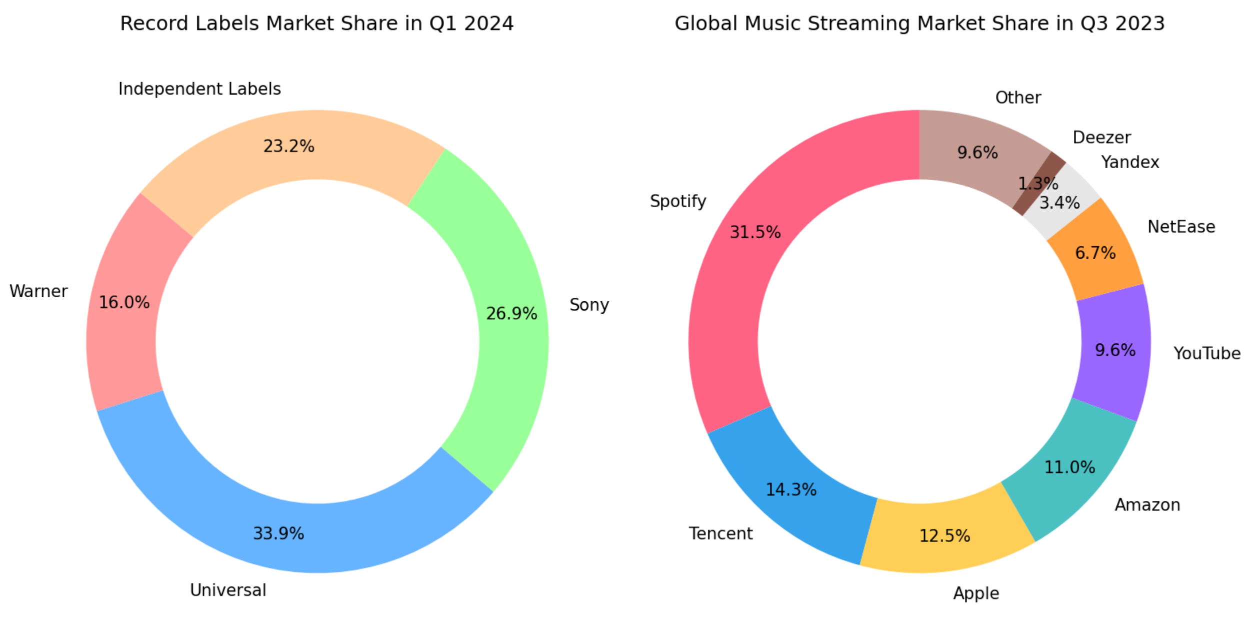

Figure 4: Concentration in the Recorded Music Industry, 2023

Streaming platforms, such as Spotify, Apple Music, and Amazon Music, are responsible 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 percent of the market. Figure 4 shows the market share of streaming platforms (and rightsholders).

These platforms are relatively undifferentiated in their music offerings, differentiating 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.[13] 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.[14]

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

14. Spotify also offers a variety of group and student subscriptions which reduce the cost per user.

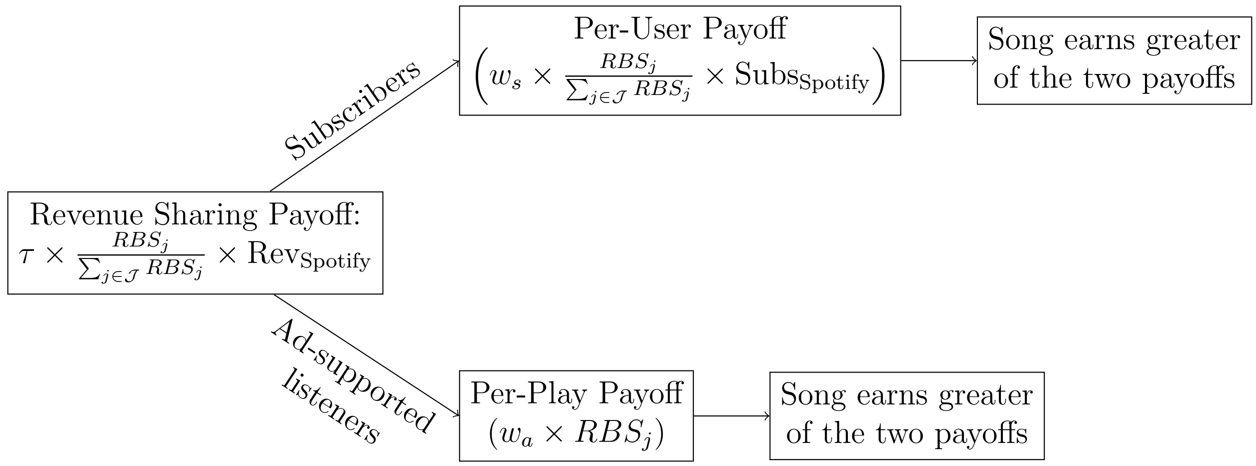

Figure 5: Revenue Sharing Payoff Structure

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.[15] Spotify pays rightsholders for royalty-bearing streams (RBS), defined as any play of a song that lasts more than 30 seconds.[16] 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.

15. Singleton 2015

16. 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

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 premium average revenue per user has been well below the standard per-subscriber fee, primarily because of family and student plans, which reduce the cost 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 percent of gross revenue.[17]

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 inventive 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.

17. Specifically, labels receive 52 percent, and publishers receive another 10-12 percent.

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 160mn 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 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 listend 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 percent 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 180mn observations, representing approximately 10 percent 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.[18] 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 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.

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

3.1 Descriptive Statistics

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

	Mean	Median	Standard Deviation	Min	Max
Duration (s)	203.27	199.32	54.28	30.13	943.53
Release Year	2019.05	2019.00	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)

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.

Most of these characteristics are uncorrelated with each other, except for loudness and

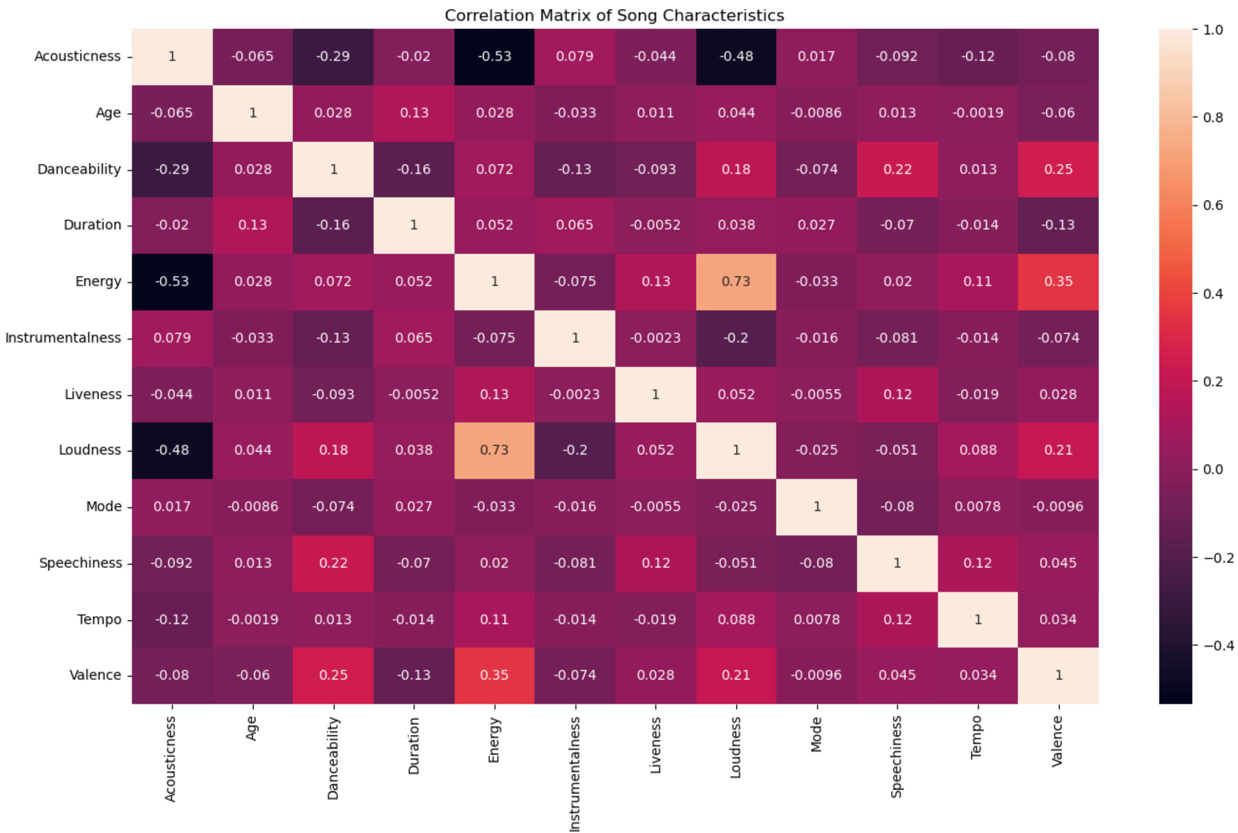

Figure 6: Correlation Matrix of Spotify Charts Song Characteristics

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:

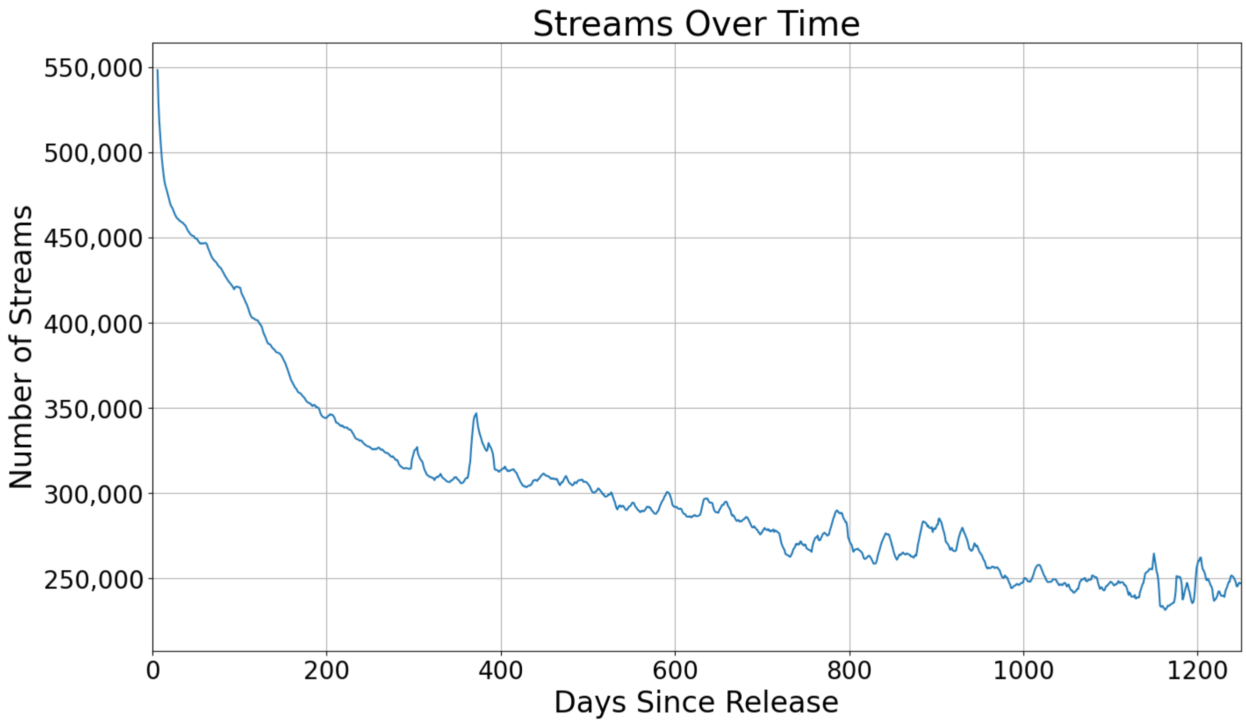

Figure 7: Number of Streams of Songs on Spotify's Top 200, by Days since Release

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 $400,000$ 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

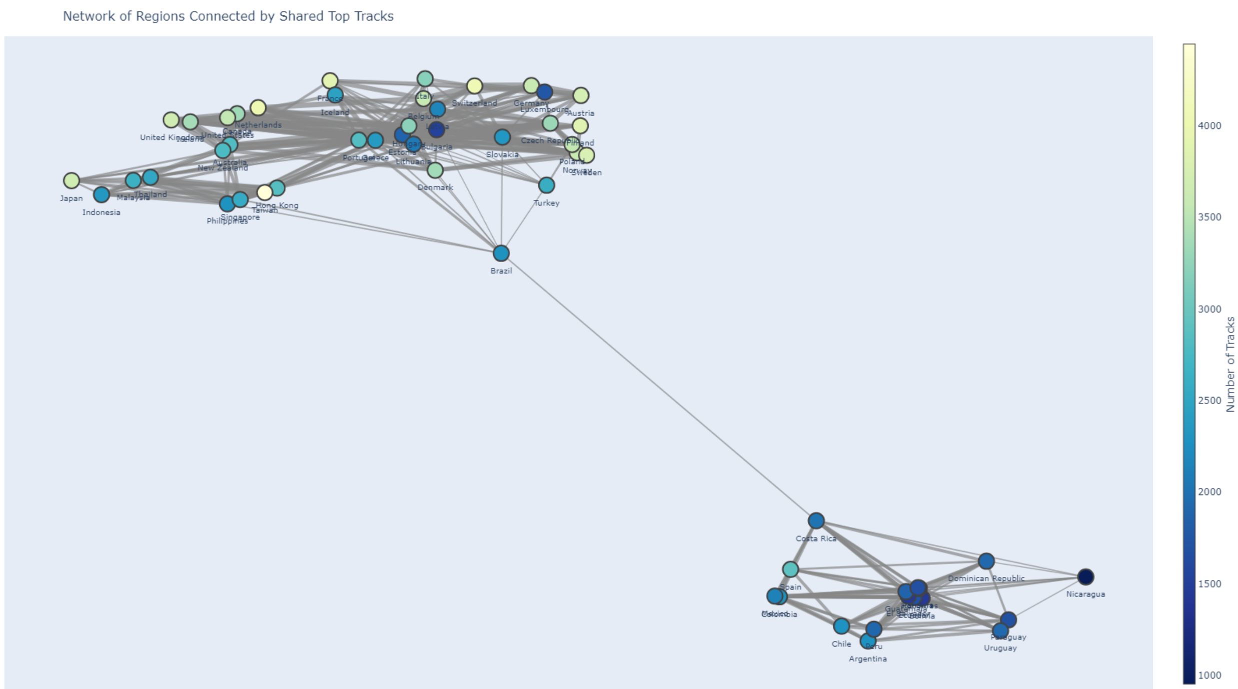

Figure 8: Spotify Charts Song Network

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

	Mean	Median	Standard Deviation	Min	Max
Duration (s)	233.19	217.91	108.40	30.00	1800.00
Release Year	2009.26	2013.00	11.03	1950.00	2019.00
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 ($N = 3.7mn$ songs)

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 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, 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.

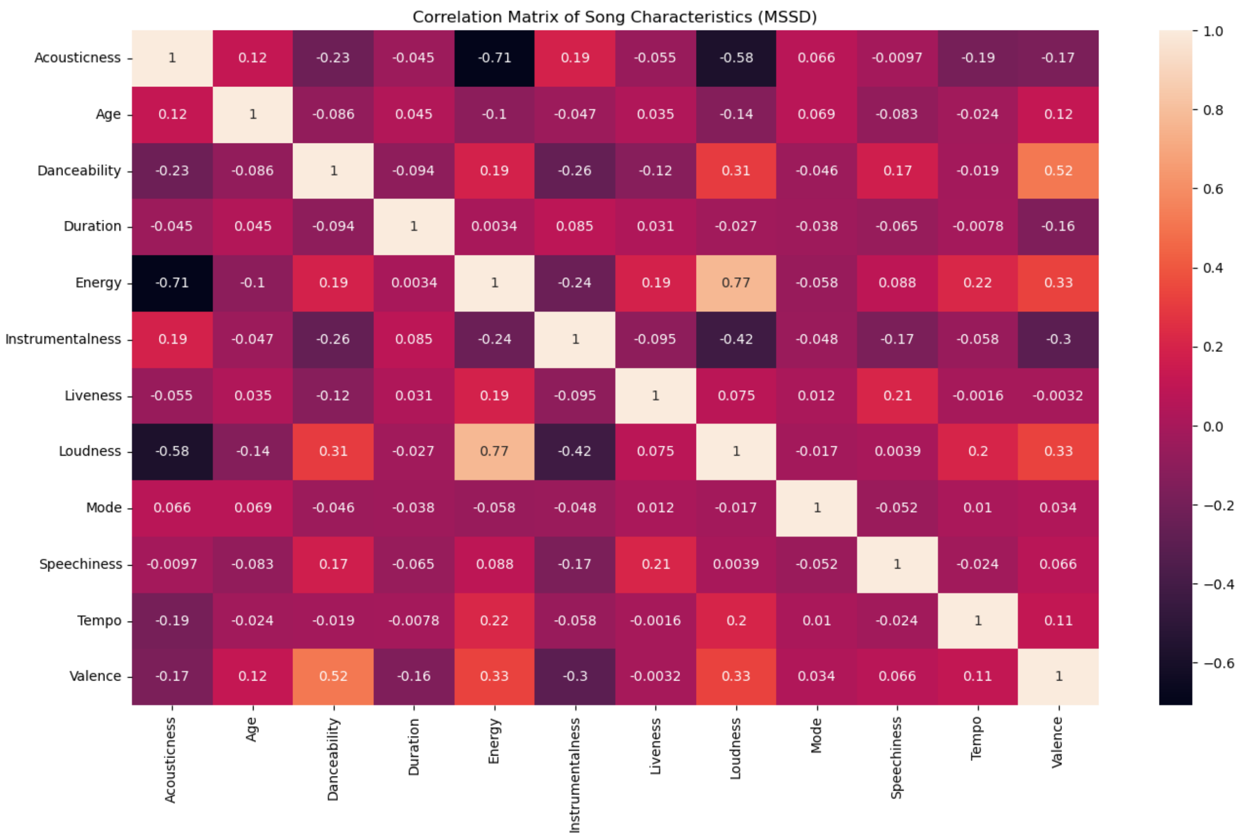

Figure 9: Correlation Matrix of MSSD Song Characteristics

	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 ($N = 180mn$song-consumer interactions)

Consumers in my sample are primarily premium subscribers, with 84 percent of the sample being premium subscribers. This is significantly higher than the percentage of premium subscribers Spotify reports, which is 40 percent of its user base.[19] It is, however, more representative of the percentage of revenue Spotify earns from premium subscribers, which is 88 percent of its revenue.[20] 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 percent of sessions being shuffle sessions. These listeners are also rather active: while 58 percent of consumer-song interactions are long enough to be considered an RBS, consumers only complete 34 percent of the songs they receive. Listening time is even throughout the week, with 13-15 percent 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 percent of sessions occurring during this time.

Consumers in my sample primarily listen to music from their own search process, or from their own collections, with 66 percent of sessions being from these sources. Algorithmic playlists and radio stations consist of 18 percent of streaming sessions. Editorial (human-curated) playlists and top charts are the least common source of music, with only 16 percent 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

19. Spotify Q2 2024 Earnings Report
20. Spotify Q2 2024 Earnings Report

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.

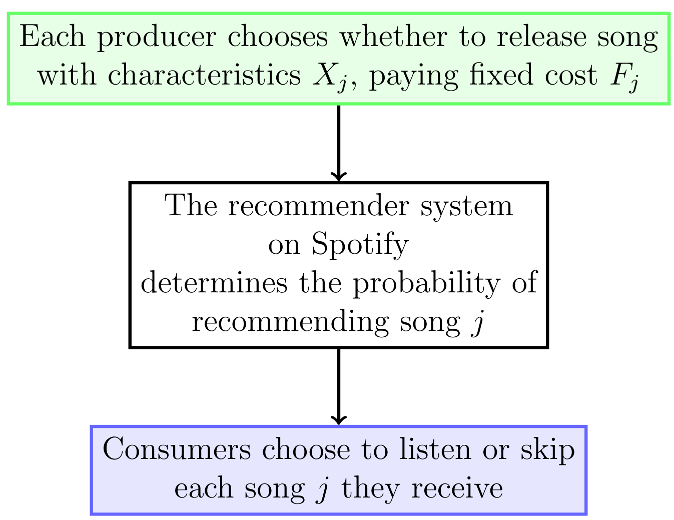

Figure 10: Timing of the Model in Each Period

4.1 Demand

Consumers in my demand model are subscribers to a streaming platform offering them a catalog of songs.[21] Each day, these consumers open the streaming app and start receiving

21. 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.

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

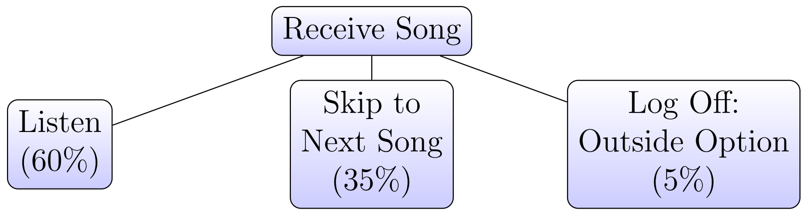

Figure 11: Consumer Decision Tree

I maintain one assumption about consumers in my model:

Assumption 1 *Consumers do not consider how their choice affects future recommendations.*[22]

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:

$$U_{L,ijs} = \beta X_j + \gamma Y_i + \eta_s + \epsilon_{ijs} \tag{2}$$

In this utility function, X_j are a vector of song characteristics and their quadratics (alternative-specific variables), Y_i are a vector of consumer characteristics (case-specific variables), η_s are position-specific fixed-effects, and ϵ_{ijs} is a Type 1 (Gumbel) Extreme Value error term. Intuitively, consumers prefer certain types of music, which I decompose in to quantitative characteristics, and their utility from a particular song may depend on when

22. 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.

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:

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

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.[23] 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.[24] 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.

23. Amazon uses collaborative filtering when recommending products "people like you also bought".

24. Continuing the Amazon example, they use content-based recommendations when describing "similar products".

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.

My model's recommender system has the following objective function:

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

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 turn to how to combine these probabilities into the demand that rightsholders face on Spotify. I define this unconditional demand as the joint probability Spotify recommends a song and a consumer listens to it:

$$
\begin{aligned}
P(i \text{ listens to } j) &= P(\text{RS surfaces } j \text{ to } i) \times P(i \text{ listens to } j | \text{RS surfaces } j \text{ to } i) \\
&= \frac{\exp(\eta_1 X_{1j} + \eta_2 \prod_{n=2}^{N} X_{nj} + \eta_3 Y_i)}{1 + \exp(\eta_1 X_{1j} + \eta_2 \prod_{n=2}^{N} X_{nj} + \eta_3 Y_i)} \\
&\times \frac{\exp(\beta X_j + \gamma Y_i + \eta_s)}{1 + (\exp(\beta X_j + \gamma Y_i + \eta_s) + \exp(\beta E_{is}[X_j | X_{j,s-1}] + \gamma Y_i + \eta_s))}
\end{aligned} \tag{7}
$$

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.

In constructing this joint probability, I assume that consumers only receive recommended songs, but this demand structure is effective for newly released songs, about whom consumers may not have ex-ante information. These are the songs whose cost I estimate in the supply side of my model.

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.[25]

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

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

rightsholders in my model:

Assumption 2 *Each song has an independent rightsholder (i.e., no multiproduct 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 her song
- The evolution of the recommender system (i.e., the probability her 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:

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

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 onetime fixed cost to release song j on Spotify, which varies by period.

I further simplify the streamshare equation by cancelling terms:

$$
\begin{aligned}
s_{jt} &= \frac{R\hat{B}S_{jt}}{\sum_k R\hat{B}S_{kt}} \\
&= \frac{\frac{\exp(\eta_1 X_{1j}+\eta_2 \prod_{n=2}^{N} X_{nj}+\eta_3 Y_i)}{1+\exp(\eta_1 X_{1j}+\eta_2 \prod_{n=2}^{N} X_{nj}+\eta_3 Y_i)} \times \frac{\exp(\beta X_j+\gamma Y_i+\eta_s)}{1+(\exp(\beta X_j+\gamma Y_i+\eta_s)+\exp(\beta E_{is}[X_j|X_{j,s-1}]+\gamma Y_i+\eta_s))}}{\sum_k \frac{\exp(\eta_1 X_{1k}+\eta_2 \prod_{n=2}^{N} X_{nk}+\eta_3 Y_i)}{1+\exp(\eta_1 X_{1k}+\eta_2 \prod_{n=2}^{N} X_{nk}+\eta_3 Y_i)} \times \frac{\exp(\beta X_k+\gamma Y_i+\eta_s)}{1+(\exp(\beta X_k+\gamma Y_i+\eta_s)+\exp(\beta E_{is}[X_k|X_{k,s-1}]+\gamma Y_i+\eta_s))}} \\
&= \frac{\frac{\exp(\eta_1 X_{1j}+\eta_2 \prod_{n=2}^{N} X_{nj}+\eta_3 Y_i)}{1+\exp(\eta_1 X_{1j}+\eta_2 \prod_{n=2}^{N} X_{nj}+\eta_3 Y_i)} \times \exp\left(\beta X_j+\gamma Y_i+\eta_s\right)}{\sum_k \frac{\exp(\eta_1 X_{1k}+\eta_2 \prod_{n=2}^{N} X_{nk}+\eta_3 Y_i)}{1+\exp(\eta_1 X_{1k}+\eta_2 \prod_{n=2}^{N} X_{nk}+\eta_3 Y_i)} \times \exp\left(\beta X_k+\gamma Y_i+\eta_s\right)}
\end{aligned}
$$

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 is key to identifying the fixed cost of releasing a song. Because firms will enter up to the breakeven point, the marginal (or worst-performing) song will just break even. That is, its expected revenue will equal the fixed cost. Similar to Aguiar and Waldfogel 2018, I employ this condition to estimate the fixed cost of releasing a song on Spotify.

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 up to the breakeven point, where the expected profit from releasing the song equals the fixed cost; and the fixed cost of releasing a song is equal to the expected revenue of the marginal (or worst-performing) song in each period.

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 calculation

In the first stage, I estimate consumer preferences and recommender system preferences using the MSSD data. 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 choice set for each consumer 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, and I apply my equilibrium condition to identify the fixed cost of releasing a song on Spotify. 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.

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 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.

I take two approaches to determine the number of songs in the streamshare measure. In the first approach, 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.[26] 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.

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

In my second 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 40mn in 2018.[27] 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. I provide results for both approaches, but I focus primarily on the first approach.

6 Results

Table 6 reports consumer demand estimates:

The demand estimates identify what characteristics consumers prefer in songs. For many of these characteristics, such as acousticness, liveness, loudness, tempo, and valence, the linear term is positive, and the quadratic term is negative. This suggests that consumers have an ideal level for these characteristics, and that songs that deviate from this ideal level are less likely to be listened to. Other characteristics, such as duration, instrumentalness, and danceability, have negative linear and quadratic terms, suggesting that consumers prefer having less of these characteristics. Consumers also prefer newer songs, with the age coefficient being negative. They also prefer songs in major mode, and with standard time signatures.

These estimates also help identify consumer characteristics that drive listening. Consumers are less likely to listen to songs in the evening, likely because of other interruptions they face in the evening hours (e.g., other commitments). Premium users are more likely to listen to songs, likely because they are not interrupted by ads.

This model also has a high $\bar{\rho}^2$, suggesting that it explains a large amount of the variation in the data. Almost all coefficients are statistically significant at the one percent level.

Table 7 reports the results for the recommender system, both with and without interaction terms:

27. Spotify 2018 Annual Report

	Dependent variable: RBS
Intercept (Listen)	3.66***
	(0.003)
Intercept (Skip)	2.71***
	(0.003)
Acousticness	0.0581***
	(0.0006)
Age	-0.0168***
	(0.00047)
Danceability	-0.00445***
	(0.00107)
Duration	-0.00204**
	(0.00081)
Duration^2	-0.00054***
	(0.000084)
Energy	-0.0122***
	(0.00126)
Instrumentalness	-0.00596***
	(0.00049)
Liveness	0.0114***
	(0.00051)
Loudness	0.00694***
	(0.00099)
Mode	0.0307***
	(0.00097)
Speechiness	-0.004***
	(0.00049)
Tempo	0.000132
	(0.00089)
Time Signature	0.0383***
	(0.00304)
Valence	0.0272***
	(0.00071)
Morning	0.144***
	(0.00042)
Afternoon	0.0834***
	(0.00037)
Night	0.13***
	(0.00061)
Premium	0.0298***
	(0.00042)
Observations	180,061,351
$\bar{\rho}^2$	0.765

Note: *p<0.1; **p<0.05; ***p<0.01

Day-of-week, session position, and quadratic terms omitted for brevity. See Appendix

Table 6: Consumer Demand Estimates

	Dependent variable: Completed Song	
	(No Interaction Terms)	(Interaction Terms)
Intercept	-0.682***	-0.660***
	(0.001)	(0.001)
Acousticness	0.043***	0.047***
	(0.000)	(0.000)
Age	-0.044***	-0.045***
	(0.000)	(0.000)
Danceability	-0.077***	-0.065***
	(0.000)	(0.001)
Duration	-0.222***	-0.218***
	(0.000)	(0.000)
Energy	0.020***	-0.019***
	(0.000)	(0.001)
Instrumentalness	0.056***	0.060***
	(0.000)	(0.000)
Loudness	-0.042***	-0.013***
	(0.000)	(0.001)
Mode	0.016***	0.014***
	(0.000)	(0.000)
Speechiness	-0.033***	-0.052***
	(0.000)	(0.000)
Time Signature	0.035***	0.012***
	(0.001)	(0.001)
Tempo	-0.011***	-0.008***
	(0.000)	(0.000)
Valence	0.036***	0.047***
	(0.000)	(0.000)
Morning	0.154***	0.154***
	(0.000)	(0.000)
Afternoon	0.097***	0.097***
	(0.000)	(0.000)
Night	0.171***	0.168***
	(0.001)	(0.001)
Premium	-0.065***	-0.069***
	(0.000)	(0.000)
Observations	180,061,351	180,061,351
Pseudo R^2	0.006	0.007

Note: *p<0.1; **p<0.05; ***p<0.01

Day-of-week, session position, and interaction terms omitted for brevity. See Appendix

Table 7: Recommender System Estimates

The recommender system estimates suggest that consumers are likelier to complete shorter, newer, and lower-tempo songs. They are also likely to complete happier (higher valence) and more instrumental songs. At the consumer level, consumers are likelier to complete songs at times other than evening, and premium users are less likely to complete songs. 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.

Introducing interactions does not materially affect many of these estimates. The only term whose sign flips is energy, which is now negative. While most of the interaction terms are statistically significant, their values are in the thousandths, suggesting that they do not have a large effect on the probability a consumer completes a song. The Pseudo R^2 also does not increase much (in absolute terms) with the introduction of interaction terms, suggesting that they do not add much explanatory power to the model. I therefore use the model without interaction terms in the rest of my analysis.

Notable differences exist between the demand and recommender system estimates. The coefficient on duration is much higher in the recommender system, suggesting that duration affects song completion more strongly than it affects RBS amounts. This also means that the recommender system is more likely to surface shorter songs than consumers may prefer. Most other characteristics have similar effects in both models, but some have different magnitudes. Age is more negative in the recommender system, possibly because consumers are hearing these newer songs for the first time. This does imply that the recommender system is more likely to surface newer songs than consumers may prefer. The only coefficients whose signs are different are energy (whose sign matches in the model with interactions), tempo (which is imprecisely estimated in the demand model), and premium status. The difference in premium status is likely because premium listeners are both more willing to try songs, but also more willing to move onto new songs without completing them.

Table 8 reports the results for the song characteristic Markov processes:

	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)
$\text{Acousticness}_{t-1}$	**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)
$\text{Danceability}_{t-1}$	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)
$\text{Instrumentalness}_{t-1}$	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)
$\text{Time Signature}_{t-1}$	-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 8: Vector Autoregression (VAR) Model Results

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 9 reports the results for the recommender system Markov process estimation:

	Dependent variable: Predicted Probability ($\hat{\phi}_t$)
$\hat{\phi}_{t-1}$	0.695***
	(0.020)
Drift	0.000***
	(0.000)
Constant	0.103***
	(0.007)
Observations	1826
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 9: Markov Process Estimation for Recommender System

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:

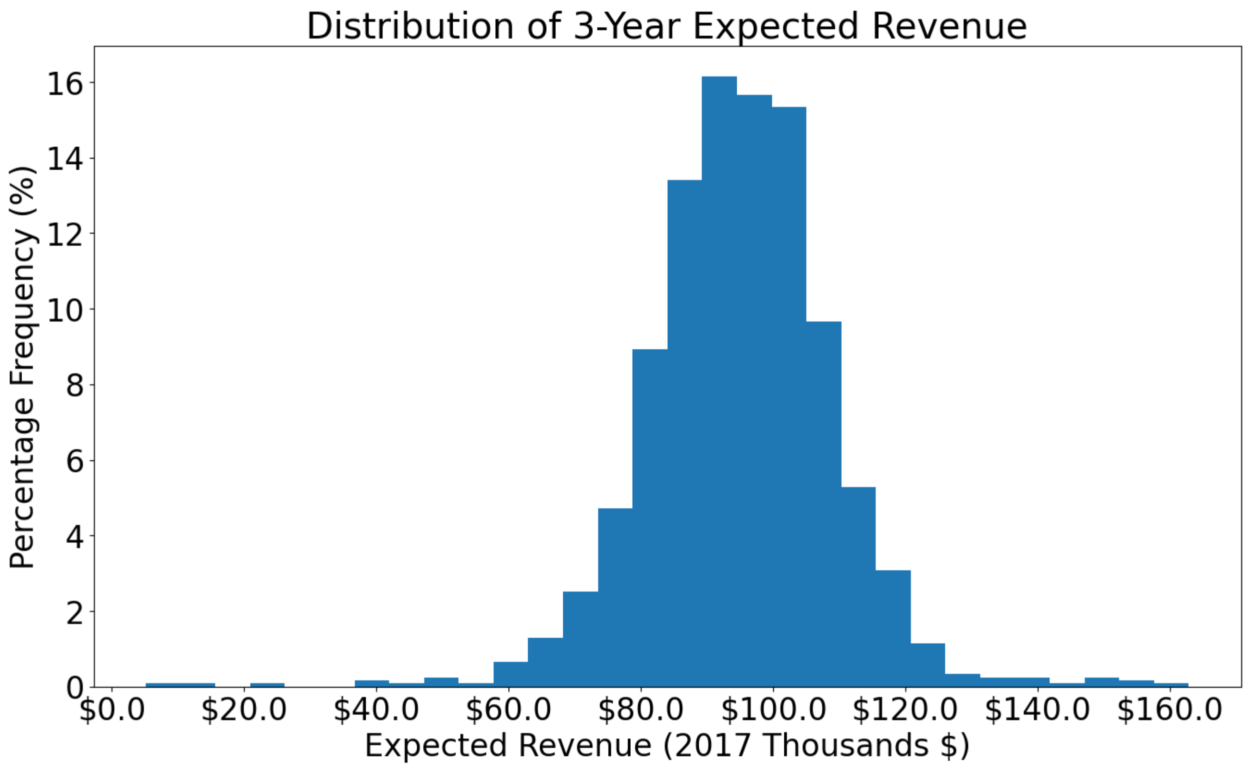

Figure 12: Expected Revenue of Songs Released in 2018 that Entered Spotify's Top 200

These songs have an expected revenue ranging from $10,000 to $160,000, with a mean at $95,000. The distribution is similar for all songs, but with much lower amounts, as the rival songs include all songs on the platform. I compute a mean expected revenue of $45.45 for all songs. Applying my equilibrium condition this data, I estimate the fixed cost by taking the expected revenue of the marginal song on any given day. Figure 13 plots the distribution of fixed costs:

The fixed cost of releasing a song on Spotify ranges from $10,000 to $135,000, with a

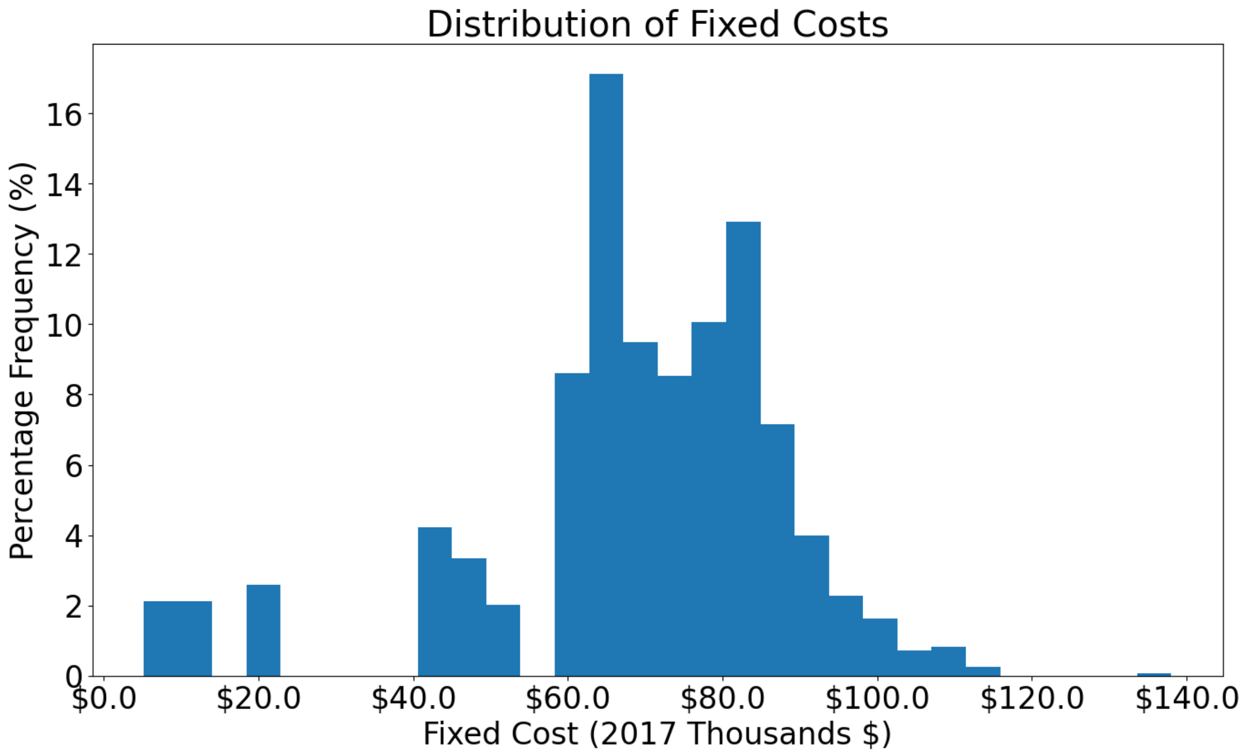

Figure 13: Fixed Costs of Songs Released in 2018 that Entered Spotify's Top 200

mean of $85,000. This distribution has a longer left-tail, with more songs clustered around the $10,000-$60,000 range. Again, the distribution is similar for all songs, but with lower amounts. I compute a mean fixed cost of $40.81 for all songs.

Figure 14 plots the distribution of fixed costs by day in 2018. They represent the unique fixed costs estimated by the model:

This distribution has the same support as the distribution of fixed costs, but is more normally distributed, with a mean of $85,000.

My estimated mean fixed cost for songs in the top 200 is close to a report from Chace 2011, which estimated the non-marketing costs of Rihanna's "Man Down" at $78,000 in 2011 dollars ($88,000 in 2017 dollars). The mean for all songs 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

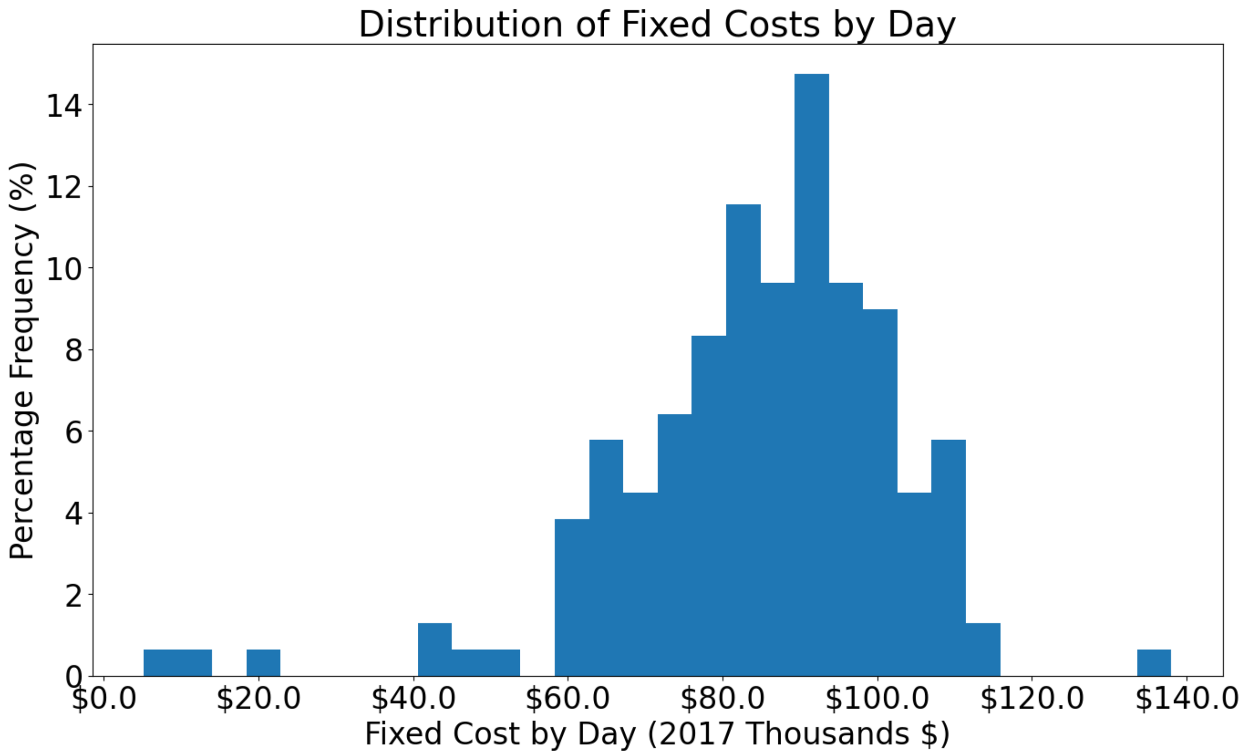

Figure 14: Distribution of Fixed Costs of Songs

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 fixed costs by day, resulting in a distribution of fixed costs. When using Aguiar and Waldfogel's of estimating the fixed cost for the entire release period, I compute a fixed cost of $3.42, which is close to their perfect foresight estimate of $6.09 ($6.79 in 2017 dollars). 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.

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 several counterfactuals. 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:

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 20 percent chance of being recommended, and compare it to the estimated revenue generated by the model.

First, I compare average expected profit for songs. Figure 15 reports the results of this comparison:

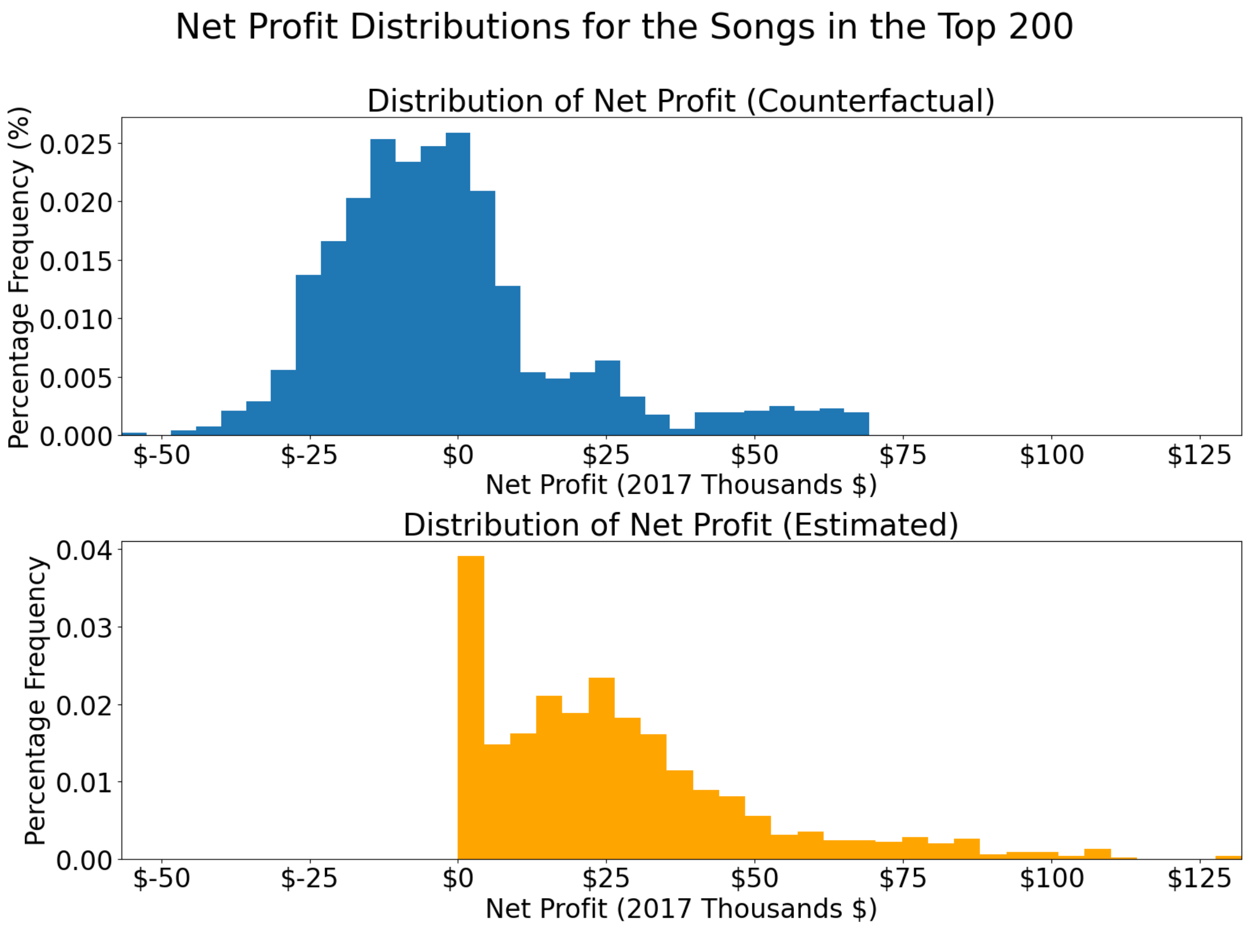

Figure 15: Counterfactual Expected Profit - Random Recommendations

Each observation in this figure represents a song released in the first three quarters of 2018. Note that the clustering at zero for estimated profit is a result of the equilibrium condition, which requires that the expected profit of the marginal song each day be zero. It is immediately apparent that many songs become unprofitable when random recommendations are used. Indeed, of the 1232 songs I observe that were released in the first three quarters of 2018, only 451 are profitable.

I now turn to some song characteristic results and welfare implications of my counterfactual analysis. Figure 16 reports the average duration of songs between profitable and unprofitable songs.

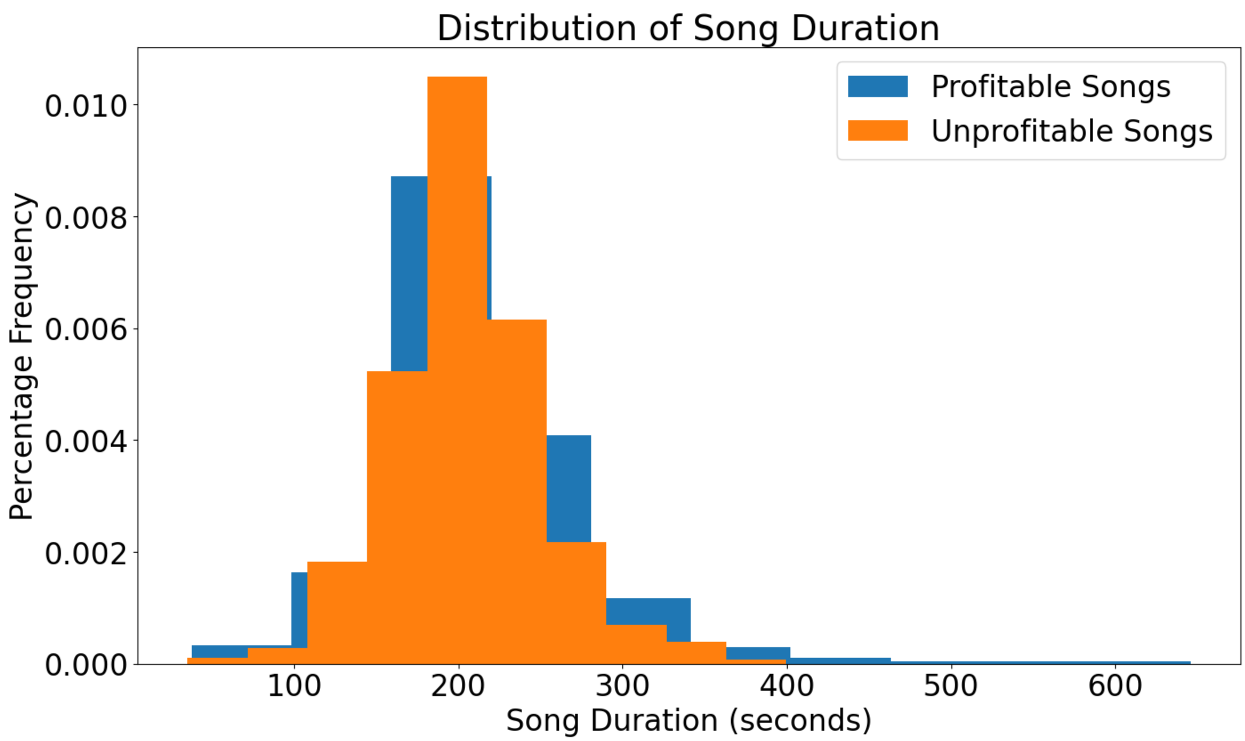

Figure 16: Counterfactual Duration - Random Recommendations

The average duration of songs of profitable songs is 213.4 seconds, and the average duration of unprofitable songs is 204.6 seconds. This difference is significant at the five percent level. Moreover, the unprofitable songs are more homogeneous, as the standard deviation of duration is 45.4 seconds, compared to 61.4 seconds for profitable songs. The difference in distributions is also significant at the five percent level. This suggests that introducing recommender systems allows shorter, more homogeneous songs to enter the market and find an audience.

Song energy represents another example of the differences between profitable and unprofitable songs. Figure 17 reports the average energy of songs between profitable and unprofitable songs.

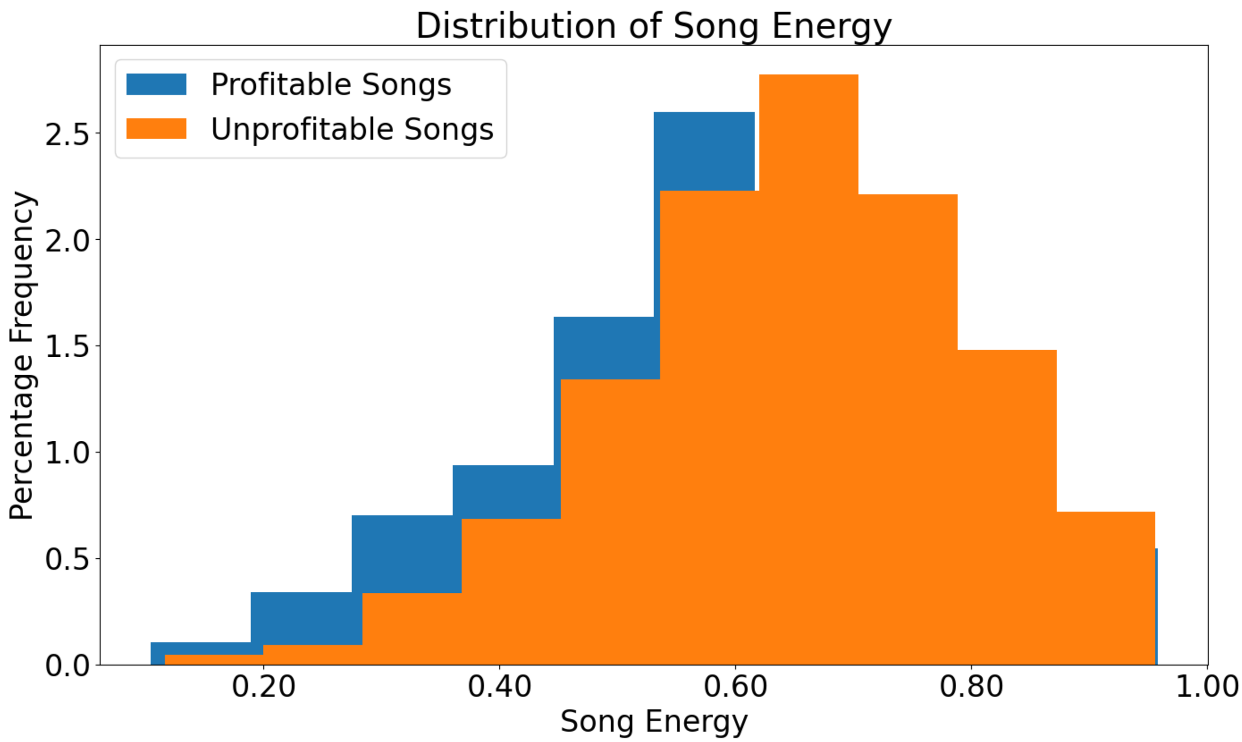

Figure 17: Counterfactual Energy - Random Recommendations

The average energy of profitable songs is 0.604, and the average energy of unprofitable songs is 0.649. This difference is significant at the one percent level. Moreover, the distribution of energy is very clearly right-shifted for unprofitable songs, compared to profitable songs.

Table 10 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):

When recommendations are random, the surviving songs are longer, less energetic, more danceable, and louder. Additionally, the distributions are different for both these variables

Feature	Profitable	Unprofitable	Diff	KS-Test P-Value
Duration (s)	213.362	204.637	-8.726**	0.0286
	(2.889)	(1.625)		
Tempo	123.845	124.584	0.739	0.6179
	(1.385)	(1.115)		
Energy	0.604	0.649	0.045***	0.0000
	(0.008)	(0.005)		
Danceability	0.712	0.688	-0.024**	0.0079
	(0.007)	(0.005)		
Valence	0.459	0.428	-0.031*	0.0001
	(0.009)	(0.008)		
Acousticness	0.222	0.190	-0.032*	0.0013
	(0.011)	(0.008)		
Instrumentalness	0.005	0.010	0.004	0.0148
	(0.002)	(0.002)		
Liveness	0.180	0.179	-0.001	0.4376
	(0.006)	(0.005)		
Speechiness	0.154	0.165	0.011	0.7059
	(0.006)	(0.005)		
Loudness	-6.875	-6.096	0.779***	0.0000
	(0.116)	(0.079)		
Mode	0.630	0.586	-0.043	0.6375
	(0.023)	(0.018)		

Table 10: Counterfactual Song Characteristics

and valence, acousticness, and instrumentalness 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 Thisse 1992. Formally, I define consumer surplus with the following equation:

$$CS = \log\left(\sum_{i=1}^{N} \exp\left(\beta X_j + \gamma Y_i + \eta_s\right)\right) \tag{12}$$

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 9.89 percent higher when targeted recommender systems are used, compared to when random recommendations are used. Restated, random recommender systems result in a 9 percent 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. These results are also consistent with the counterfactual estimates when I assume that songs are competing against all other songs on the platform, rather than just the top 200 songs. With that assumption, I find that consumer surplus is 10.3 percent higher when targeted recommender systems are used.

7.2 Oracular Recommendations

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

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 18 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 5.3 percent compared to the simulated streaming sessions. This difference is statistically significant at the one percent level.

Table 11 reports the average values of song characteristics for the simulated streaming sessions and the utility-maximizing streaming sessions, and the difference in means:

The optimal streaming session is more acoustic, less 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	0.64	-0.44***
			(-8.61)
Age	0.05	0.15	-0.10
			(-1.42)
Danceability	0.70	0.59	0.11***
			(3.42)
Duration	207.74	196.61	11.13
			(0.96)
Energy	0.63	0.27	0.36***
			(10.37)
Instrumentalness	0.01	0.00	0.01
			(0.54)
Liveness	0.18	0.22	-0.04
			(-1.37)
Loudness	-6.39	-10.82	4.43***
			(8.54)
Speechiness	0.16	0.09	0.07**
			(2.28)
Tempo	124.33	113.78	10.55
			(1.57)
Valence	0.44	0.42	0.02
			(0.38)
Consumer Surplus	3.01	3.18	-0.17***
			(-3.67)
Note:	T-statistics in parentheses; *p<0.1; **p<0.05; ***p<0.01		

Table 11: Counterfactual Song Characteristics and Consumer Surplus - Oracular Recommendations

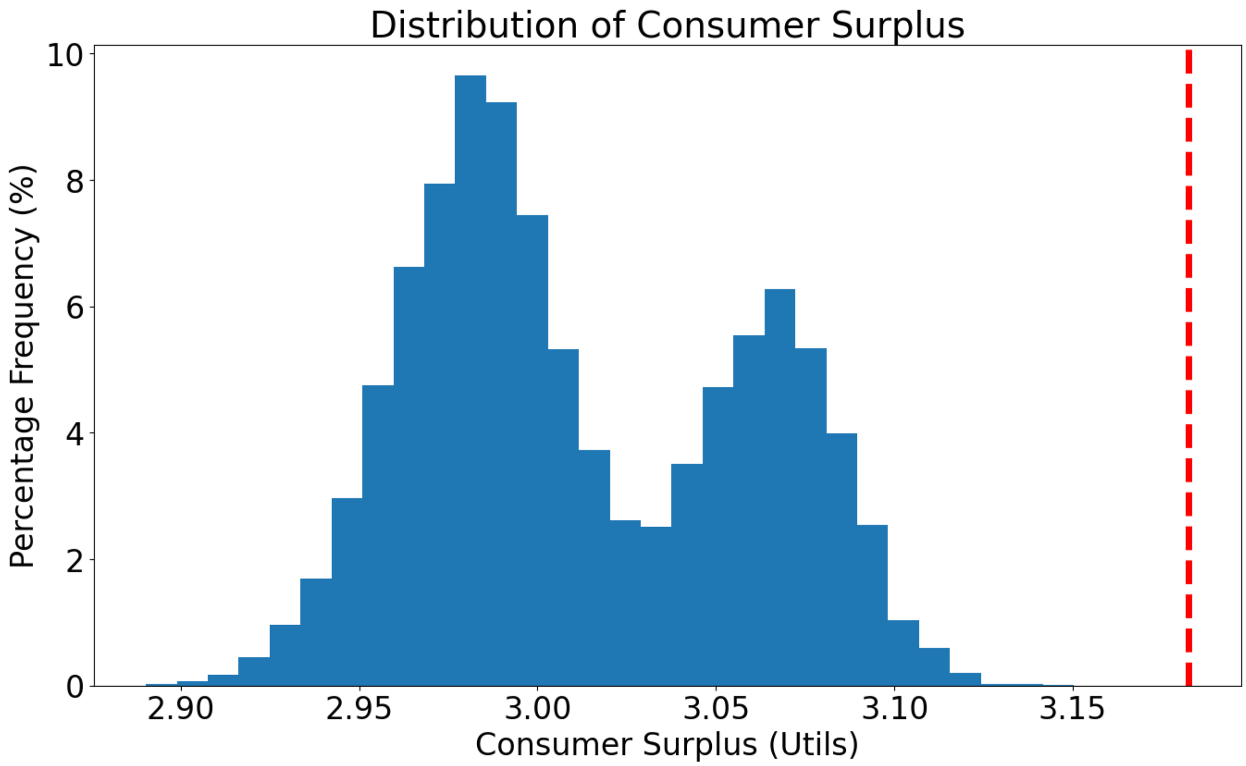

Figure 18: Counterfactual Consumer Surplus - Oracular Recommendations

7.3 Popular Recommendations

The third 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 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.

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 19 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, popular recommendations help some songs and hurt others. On average, however, songs are worse off when popular recommendations are used. Indeed, of the 1232 songs I observe that were released in the first three quarters of 2018, only 212 are profitable.

I now turn to some song characteristic results and welfare implications of my counterfactual analysis. Figure 20 reports the average duration of songs between profitable and unprofitable songs.

The average duration of songs of profitable songs is 204.6 seconds, and the average duration of unprofitable songs is 208.5 seconds. This difference is not significant at the five percent level.

Table 12 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

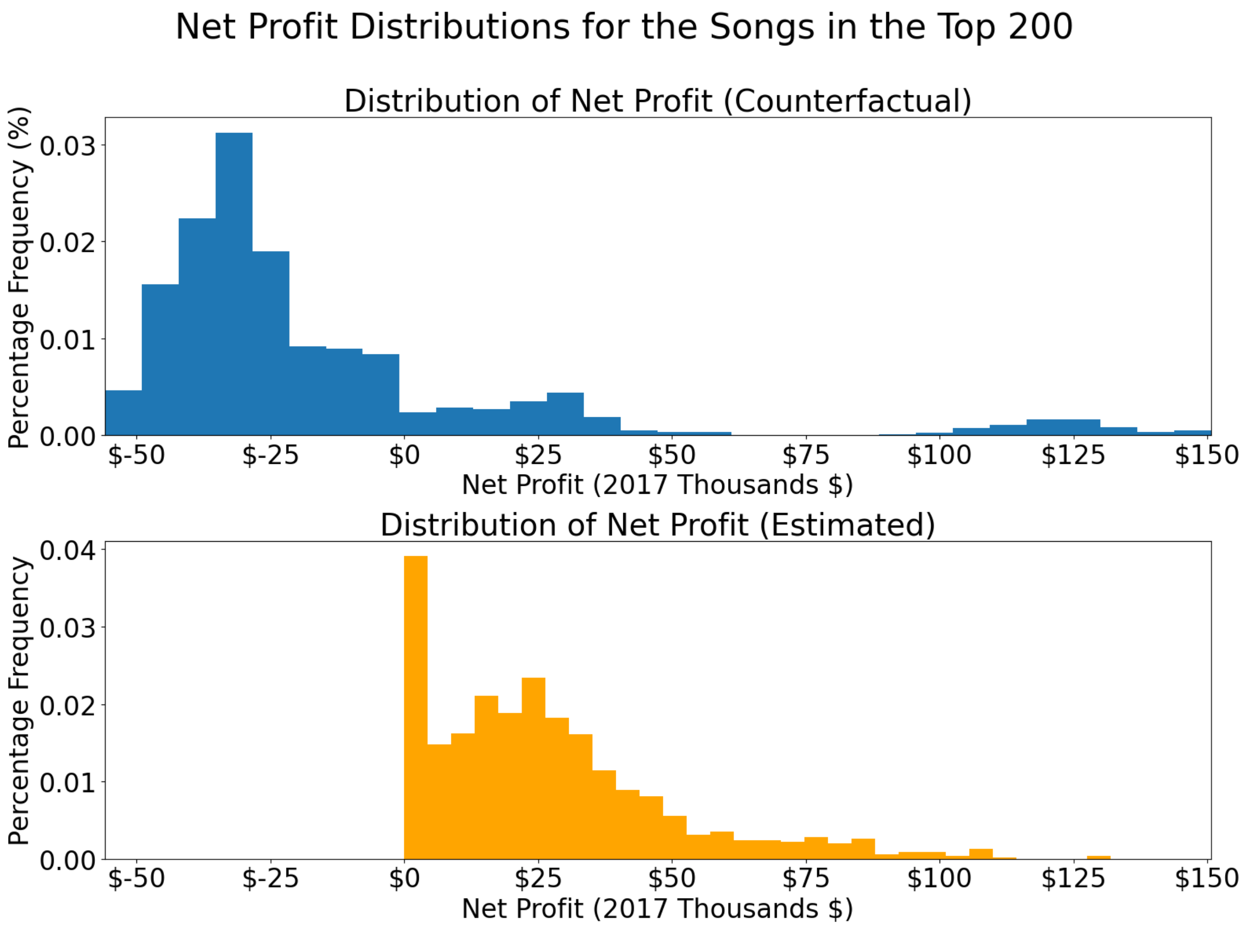

Figure 19: Counterfactual Expected Profit - Popular Recommendations

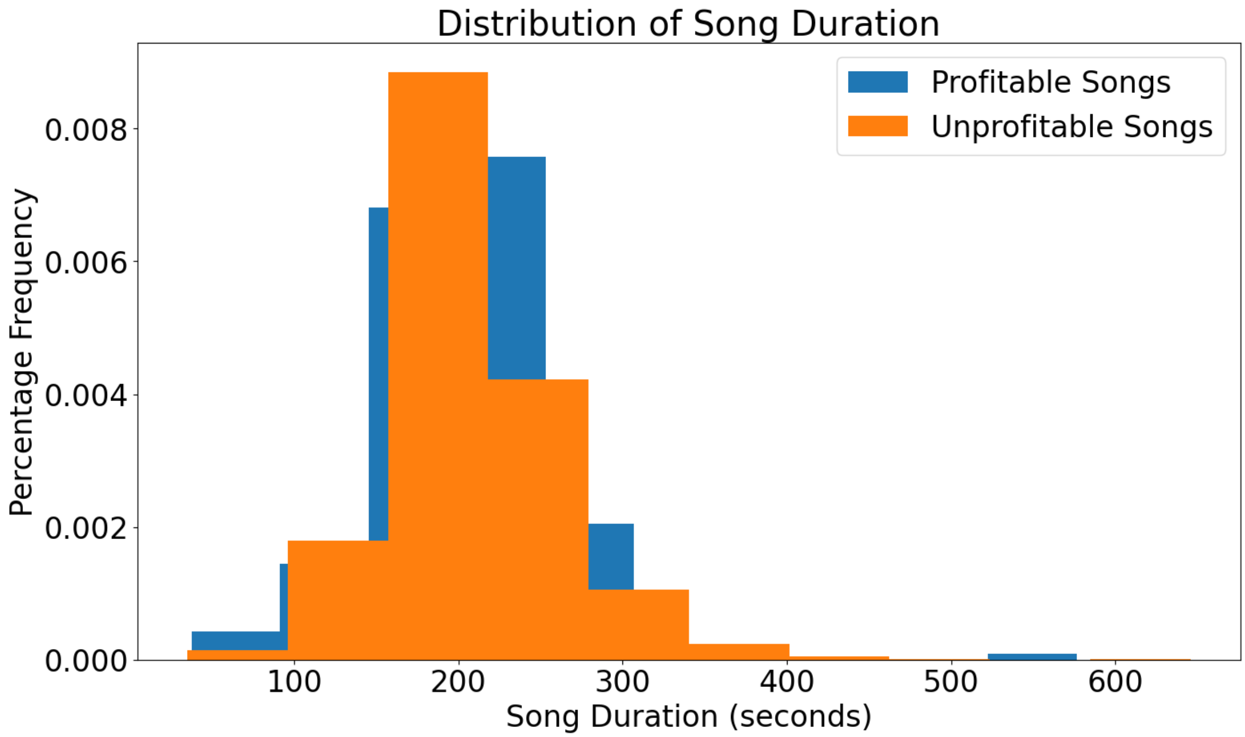

Figure 20: Counterfactual Duration - Popular Recommendations

a KS-Test):

None of the mean differences are significant at the five percent level, but some of the distributions are significantly different. At the ten percent level, the distributions of tempo and danceability are different, and at the five percent level, the distribution of valence is different.

Figure 21 plots the distribution of song valence for profitable and unprofitable songs:

While the mean of this distribution is not significantly different, the distribution is left-shifted for unprofitable songs. This suggests that in the absence of targeted recommendations, songs with lower valence (i.e., less happy-sounding songs) struggle to find their audience.

Finally, I turn to the welfare implications of my counterfactual analysis. I find that consumer surplus is 18.6 percent higher when targeted recommender systems are used, compared to popular recommendations. Restated, random recommender systems result in a 15.8 percent decrease in consumer surplus. These results are also consistent with the counterfactual

Feature	Profitable	Unprofitable	Diff	KS-Test P-Value
Duration (s)	204.607	208.524	3.916	0.5154
	(3.420)	(1.642)		
Tempo	123.542	124.479	0.938	0.0764
	(1.871)	(0.978)		
Energy	0.620	0.635	0.015	0.6203
	(0.011)	(0.005)		
Danceability	0.708	0.694	-0.014	0.0932
	(0.010)	(0.004)		
Valence	0.452	0.436	-0.016	0.0159
	(0.013)	(0.007)		
Acousticness	0.195	0.203	0.008	0.7148
	(0.015)	(0.007)		
Instrumentalness	0.006	0.008	0.003	0.8448
	(0.002)	(0.002)		
Liveness	0.166	0.182	0.017	0.3592
	(0.007)	(0.004)		
Speechiness	0.151	0.163	0.012	0.1519
	(0.008)	(0.004)		
Loudness	-6.551	-6.345	0.206	0.6057
	(0.175)	(0.072)		
Mode	0.651	0.592	-0.060	0.5230
	(0.032)	(0.015)		

Table 12: Counterfactual Song Characteristics - Popular Recommendations

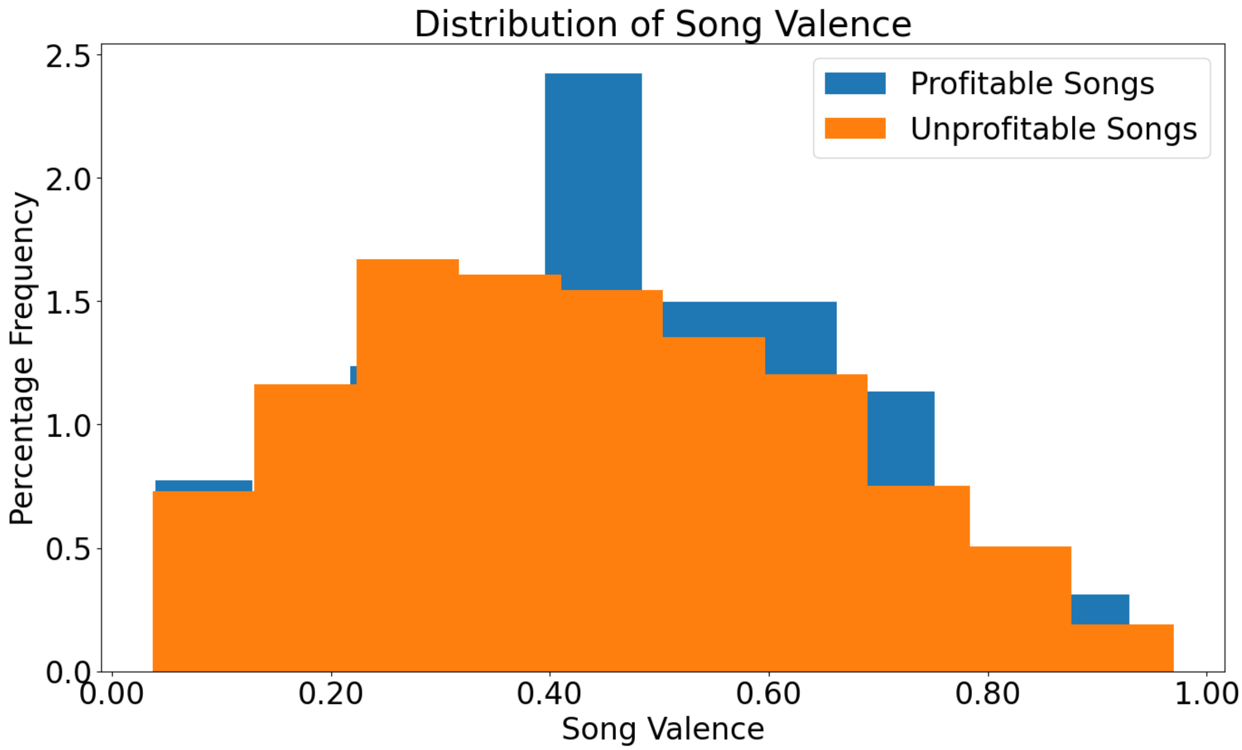

Figure 21: Counterfactual Valence - Popular Recommendations

estimates when I assume that songs are competing against all other songs on the platform, rather than just the top 200 songs. With that assumption, I find that consumer surplus is 19.1 percent higher when targeted recommender systems are used.

8 Conclusion

As recommender systems become more integrated into the US economy, it is paramount to understand the impact these systems have on both consumer demand and equilibrium supply decisions. This paper provides a structural model of the music streaming industry to estimate the impact these systems have had on music releases. Because the music industry is typically the vanguard for technology adoption, understanding the impact of recommender systems on music releases can provide insights into the impact of recommender systems on other industries, such as film, television, and shopping.

I find that recommender systems have indeed changed the sound of music, resulting in shorter, more homogeneous songs. These systems, however, allow for more songs to enter the market, earn more revenue, and increase consumer surplus approximately 9% compared to a random recommender system. I also find that the recommender system's preferences differ from consumer preferences, in part because the platform's incentives differ from rightsholders: rightsholders want consumers to listen to at least 30 seconds of their songs to earn royalties, but Spotify's recommender system rewards complete listens, to reduce the amount of royalty payouts. In the absence of recommender systems, songs are more aligned with consumer preferences, though fewer of them are profitable.

Further avenues for research exist, particularly in applying this model to other industries. The skip/listen binary choice makes this model particularly applicable to short-form video services (e.g., TikTok, Instagram Reels, YouTube Shorts), whose content is almost entirely recommender-drivel. Additionally, random coefficients would particularly enrich the consumer demand structure, and better capture the variety of preferences consumers have for music. Moreover, this model excludes the extensive margin: whether to subscribe or not. Incorporating an endogenous platform decision would better capture Spotify's pricing incentives, and how much it can trade off pricing power with the music it provides. Such an analysis would also allow for a more nuanced antitrust analysis, seeing how much recommender systems can facilitate market power.

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:

$$\begin{aligned} log(\text{Streams}_{jt}) = \alpha + \beta_1 \text{Duration}_j + \beta_2 \text{Time Trend}_t \\ + \delta(\text{Duration} \times \text{Time Trend})_{jt} + \gamma X_j + \eta_t + \epsilon_{jt} \end{aligned} \tag{13}$$

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 13 reports the results of this regression:

I find that the coefficient on the interaction is positive and significant at the one percent

	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 13: Difference-in-Differences Results

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 percent increase in streams.

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

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

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 seasonality in music listening.

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

Table 14 reports the results of this regression:

	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 14: Discrete Choice Model Results

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 most 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:

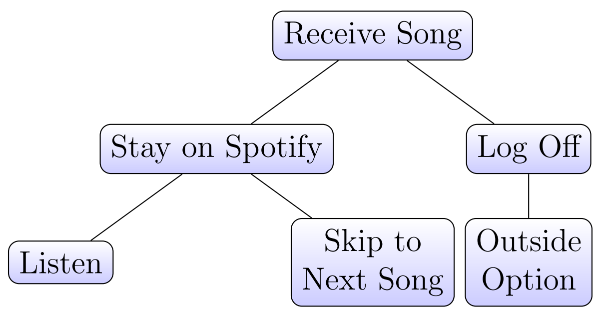

Figure 22: Nested Consumer Decision Tree

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, estimating the inside options first, then the nest parameter. Table 15 reports the results of this estimation:

	Dependent variable: Revenue Bearing Streams	
	(Nested Logit)	(Conditional Logit)
σ	1.000^{***}	
	(0.00535)	–
Constant (Listen)	3.140^{***}	3.140^{***}
	(0.0195)	(0.00128)
Constant (Skip)	2.760^{***}	2.760^{***}
	(0.0194)	(0.000503)
Acousticness	0.00740^{***}	0.00740^{***}
	(0.000252)	(0.000243)
Acousticness^2	-0.0344^{***}	-0.0344^{***}
	(0.000127)	(0.000126)
Danceability	0.0134^{***}	0.0134^{***}
	(0.000459)	(0.000447)
Danceability^2	-0.00995^{***}	-0.0100^{***}
	(0.000459)	(0.000459)
Duration	-0.00409^{***}	-0.00409^{***}
	(0.000347)	(0.000343)
Duration^2	-0.00200^{***}	-0.00200^{***}
	(9.52e-05)	(9.47e-05)
Energy	0.0420^{***}	0.0420^{***}
	(0.000439)	(0.000427)
Energy^2	0.00688^{***}	0.00686^{***}
	(0.000393)	(0.000391)

Instrumentalness	-0.0891***	-0.0891***
	(0.000382)	(0.000382)
Instrumentalness^2	0.00916***	0.00916***
	(8.42e-05)	(8.37e-05)
Liveness	-0.00244***	-0.00244***
	(0.000247)	(0.000247)
Liveness^2	-0.00672***	-0.00672***
	(8.80e-05)	(8.66e-05)
Loudness	-0.123***	-0.123***
	(0.000267)	(0.000266)
Loudness^2	-0.0177***	-0.0177***
	(0.000110)	(0.000104)
Mode	0.0249***	0.0249***
	(0.000315)	(0.000313)
Speechiness	0.0418***	0.0418***
	(0.000281)	(0.000243)
Speechiness^2	-0.0210***	-0.0210***
	(9.07e-05)	(7.28e-05)
Tempo	0.0177***	0.0176***
	(0.000311)	(0.000310)
Tempo^2	-0.0417***	-0.0417***
	(0.000467)	(0.000467)
Time Signature	-0.0170***	-0.0161***
	(0.00104)	(0.00104)
Valence	-0.00161***	-0.00161***
	(0.000278)	(0.000272)
Valence^2	-0.0615***	-0.0615***

	(0.000245)	(0.000243)
Age	-0.0283***	-0.0283***
	(0.000225)	(0.000214)
Morning	0.148***	0.148***
	(0.000423)	(0.000419)
Afternoon	0.0866***	0.0866***
	(0.000372)	(0.000372)
Night	0.119***	0.119***
	(0.000606)	(0.000604)
Premium	0.0112***	0.0112***
	(0.000415)	(0.000415)
Tuesday	-0.0143***	-0.0143***
	(0.000550)	(0.000550)
Wednesday	-0.0207***	-0.0207***
	(0.000562)	(0.000561)
Thursday	-0.0280***	-0.0280***
	(0.000559)	(0.000559)
Friday	0.000688	0.000687
	(0.000549)	(0.000549)
Saturday	-0.0221***	-0.0221***
	(0.000572)	(0.000572)
Sunday	-0.0177***	-0.0178***
	(0.000569)	(0.000569)
Observations	180,061,351	180,061,351
$\bar{\rho}^2$	0.558	0.791

Note: *p<0.1; **p<0.05; ***p<0.01

Table 15: Logit and Nested Logit Comparison

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.

Tables

	Dependent variable: Royalty-Bearing Stream
Constant (Listen)	3.660***
	(0.003)
Constant (Skip)	2.710***
	(0.003)
Acousticness	0.0581***
	(0.0006)
Acousticness^2	-0.0436***
	(0.0001)
Age	-0.0168***
	(0.00047)
Danceability	-0.00445***
	(0.00107)
Danceability^2	-0.00147***
	(0.00044)
Duration	-0.00204**
	(0.00081)
Duration^2	-0.00054***
	(8.38e-05)
Energy	-0.0122***
	(0.00126)
Energy^2	0.0603***
	(0.00037)
Instrumentalness	-0.00596***
	(0.00049)
Instrumentalness^2	-0.00958***
	(4.37e-05)
Liveness	0.0114***
	(0.00051)
Liveness^2	-0.00673***
	(6.24e-05)
Loudness	0.00694***
	(0.00099)

Loudness^2	-0.00785***
	(9.91e-05)
Mode	0.0307***
	(0.00097)
Speechiness	-0.00400***
	(0.00049)
Speechiness^2	-0.0105***
	(5.61e-05)
Tempo	0.000132
	(0.00089)
Tempo^2	-0.00799***
	(0.00044)
Time Signature	0.0383***
	(0.00304)
Valence	0.0272***
	(0.00071)
Valence^2	-0.0589***
	(0.00024)
Session Position 2	-0.333***
	(0.00096)
Session Position 3	-0.497***
	(0.00095)
Session Position 4	-0.595***
	(0.00094)
Session Position 5	-0.644***
	(0.00094)
Session Position 6	-0.677***
	(0.00094)
Session Position 7	-0.689***
	(0.00094)
Session Position 8	-0.692***
	(0.00094)
Session Position 9	-0.677***
	(0.00094)
Session Position 10	-0.637***
	(0.00094)

Session Position 11	-0.651***
	(0.00095)
Session Position 12	-0.669***
	(0.00097)
Session Position 13	-0.684***
	(0.00098)
Session Position 14	-0.698***
	(0.00100)
Session Position 15	-0.710***
	(0.00101)
Session Position 16	-0.723***
	(0.00102)
Session Position 17	-0.733***
	(0.00104)
Session Position 18	-0.741***
	(0.00106)
Session Position 19	-0.747***
	(0.00107)
Session Position 20	-0.745***
	(0.00109)
Morning	0.144***
	(0.00042)
Afternoon	0.0834***
	(0.00037)
Night	0.130***
	(0.00061)
Premium	0.0298***
	(0.00042)
Tuesday	-0.0150***
	(0.00055)
Wednesday	-0.0213***
	(0.00056)
Thursday	-0.0285***
	(0.00056)
Friday	-0.00139**

	(0.00055)
Saturday	-0.0240***
	(0.00057)
Sunday	-0.0176***
	(0.00057)
Observations	180,061,351
$\bar{\rho}^2$	0.765

Note: *p<0.1; **p<0.05; ***p<0.01

Table 17: Full Demand Estimates

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 16: Descriptions of Song Characteristics

	Dependent variable: Song Completion
	Coefficient
Intercept	-0.682***
	(0.001)
Acousticness	0.043***
	(0.000)
Danceability	-0.077***
	(0.000)
Duration	-0.222***
	(0.000)
Energy	0.020***
	(0.000)
Instrumentalness	0.056***
	(0.000)
Loudness	-0.042***
	(0.000)
Mode	0.016***
	(0.000)
Speechiness	-0.033***
	(0.000)
Tempo	-0.011***
	(0.000)
Valence	0.036***
	(0.000)
Age	-0.044***
	(0.000)
Time Signature	0.035***
	(0.001)
Session Position 2	0.038***
	(0.001)
Session Position 3	-0.019***
	(0.001)
Session Position 4	-0.073***

	(0.001)
Session Position 5	-0.090***
	(0.001)
Session Position 6	-0.107***
	(0.001)
Session Position 7	-0.104***
	(0.001)
Session Position 8	-0.097***
	(0.001)
Session Position 9	-0.066***
	(0.001)
Session Position 10	-0.015***
	(0.001)
Session Position 11	-0.022***
	(0.001)
Session Position 12	-0.034***
	(0.001)
Session Position 13	-0.040***
	(0.001)
Session Position 14	-0.049***
	(0.001)
Session Position 15	-0.056***
	(0.001)
Session Position 16	-0.064***
	(0.001)
Session Position 17	-0.068***
	(0.001)
Session Position 18	-0.073***
	(0.001)
Session Position 19	-0.074***
	(0.001)
Session Position 20	-0.064***
	(0.001)
Morning	0.154***
	(0.000)
Afternoon	0.097***

	(0.000)
Night	0.171***
	(0.001)
Premium	-0.065***
	(0.000)
Monday	(base)
Tuesday	-0.008***
	(0.001)
Wednesday	-0.009***
	(0.001)
Thursday	-0.018***
	(0.001)
Friday	-0.021***
	(0.001)
Saturday	-0.017***
	(0.001)
Sunday	-0.006***
	(0.001)
Observations	180,061,351
Pseudo R^2	0.006

Note: *p<0.1; **p<0.05; ***p<0.01

Table 18: Full Recommender System Estimates

Feature	Profitable	Unprofitable	Diff	KS-Test P-Value
Duration (s)	213.132	204.968	-8.164**	0.1244
	(2.988)	(1.604)		
Tempo	123.631	124.682	1.050	0.3738
	(1.385)	(1.112)		
Energy	0.600	0.650	0.050***	0.0000
	(0.008)	(0.005)		
Danceability	0.713	0.688	-0.025**	0.0074
	(0.007)	(0.005)		
Valence	0.457	0.430	-0.027*	0.0002
	(0.009)	(0.008)		
Acousticness	0.234	0.185	-0.049***	0.0000
	(0.011)	(0.008)		
Instrumentalness	0.005	0.010	0.005	0.0174
	(0.001)	(0.002)		
Liveness	0.179	0.179	0.000	0.8084
	(0.006)	(0.005)		
Speechiness	0.155	0.164	0.009	0.6203
	(0.006)	(0.005)		
Loudness	-6.931	-6.084	0.847***	0.0000
	(0.119)	(0.078)		
Mode	0.644	0.580	-0.064*	0.1972
	(0.023)	(0.017)		

Table 19: Counterfactual Song Characteristics for Random Recommendations - Alternate Streamshare Measure

Feature	Profitable	Unprofitable	Diff	KS-Test P-Value
Duration (s)	205.434	208.329	2.895	0.3813
	(3.447)	(1.639)		
Tempo	123.899	124.399	0.501	0.0955
	(1.891)	(0.975)		
Energy	0.621	0.635	0.014	0.5716
	(0.011)	(0.005)		
Danceability	0.709	0.694	-0.015	0.1166
	(0.010)	(0.004)		
Valence	0.456	0.436	-0.021	0.0046
	(0.013)	(0.007)		
Acousticness	0.190	0.205	0.015	0.4172
	(0.015)	(0.007)		
Instrumentalness	0.006	0.008	0.003	0.9815
	(0.002)	(0.002)		
Liveness	0.165	0.182	0.017	0.2049
	(0.008)	(0.004)		
Speechiness	0.151	0.163	0.012	0.1139
	(0.009)	(0.004)		
Loudness	-6.516	-6.353	0.163	0.5315
	(0.175)	(0.072)		
Mode	0.642	0.594	-0.047	0.8047
	(0.033)	(0.015)		

Table 20: Counterfactual Song Characteristics for Popular Recommendations - Alternate Streamshare Measure

Figures

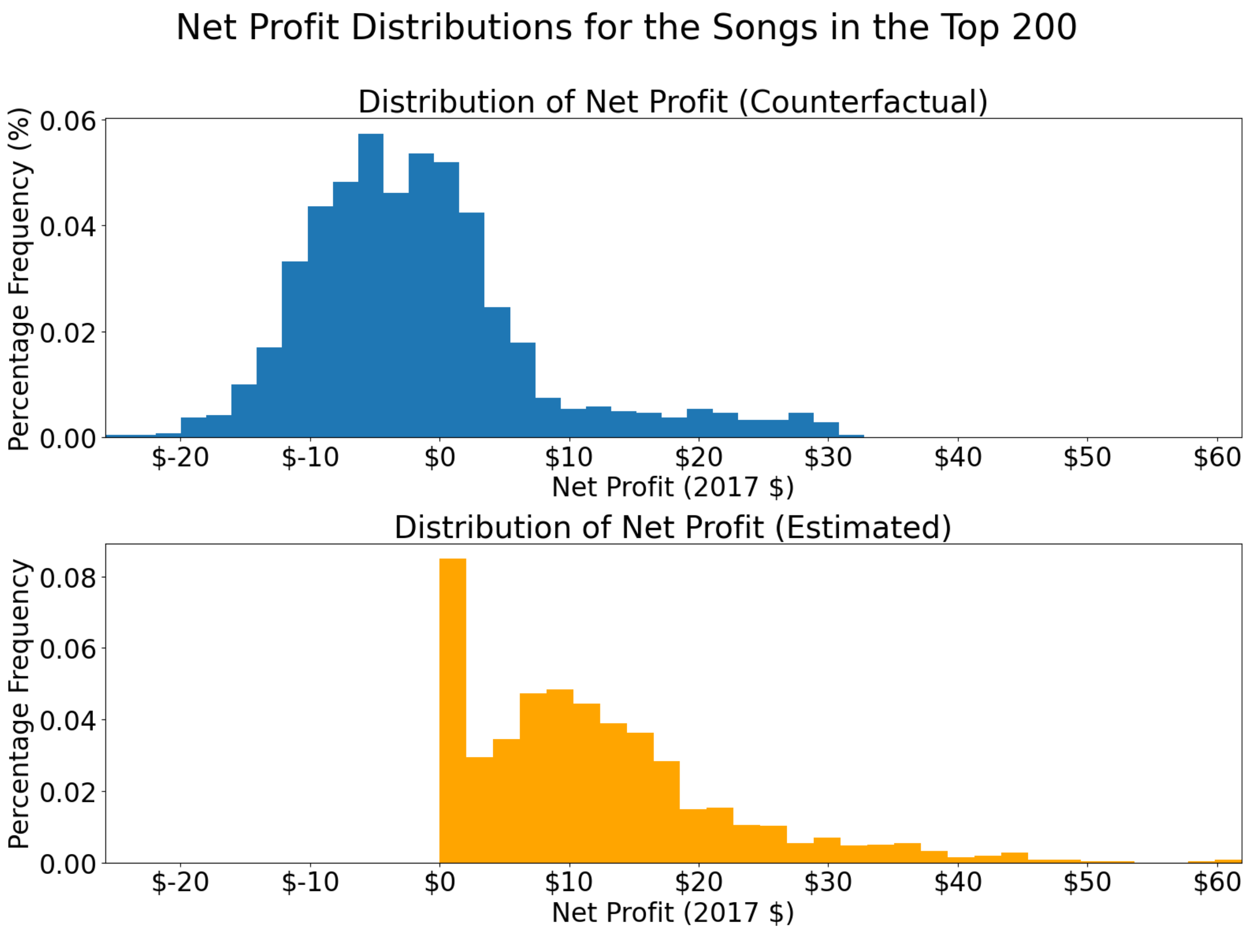

Figure 23: Counterfactual Expected Profit - Random Recommendations

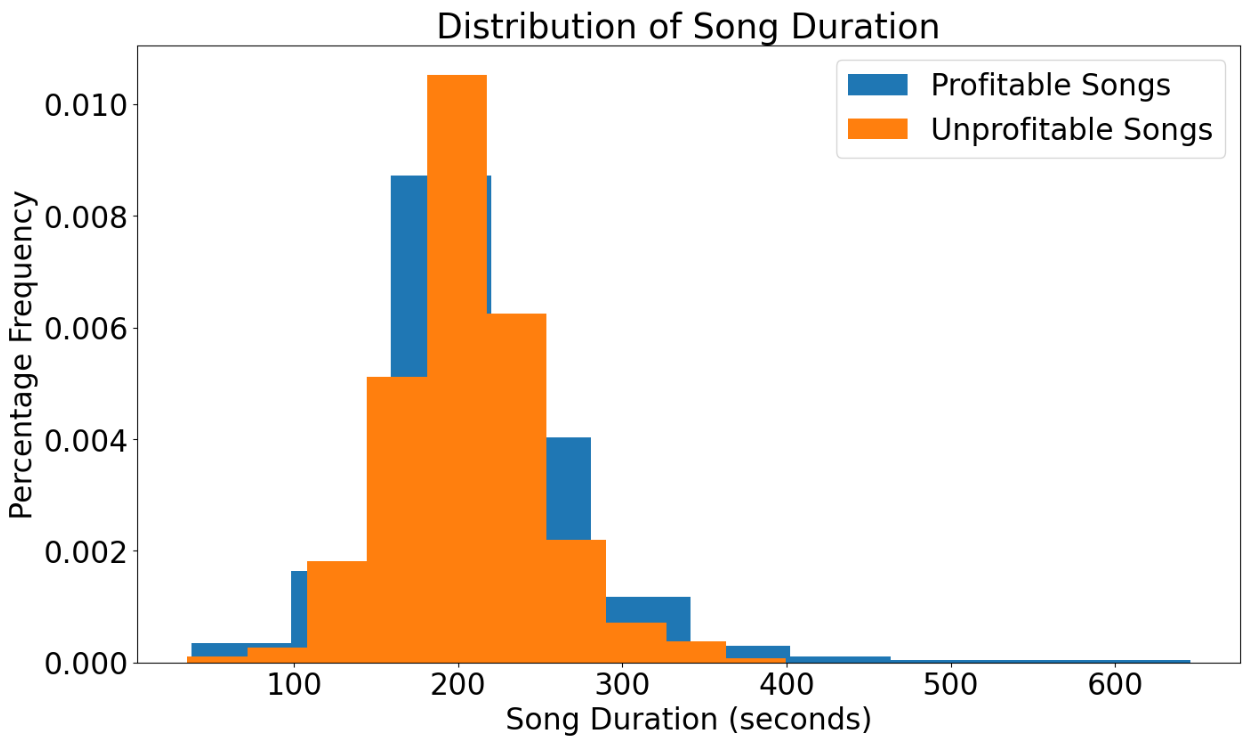

Figure 24: Counterfactual Duration - Random Recommendations

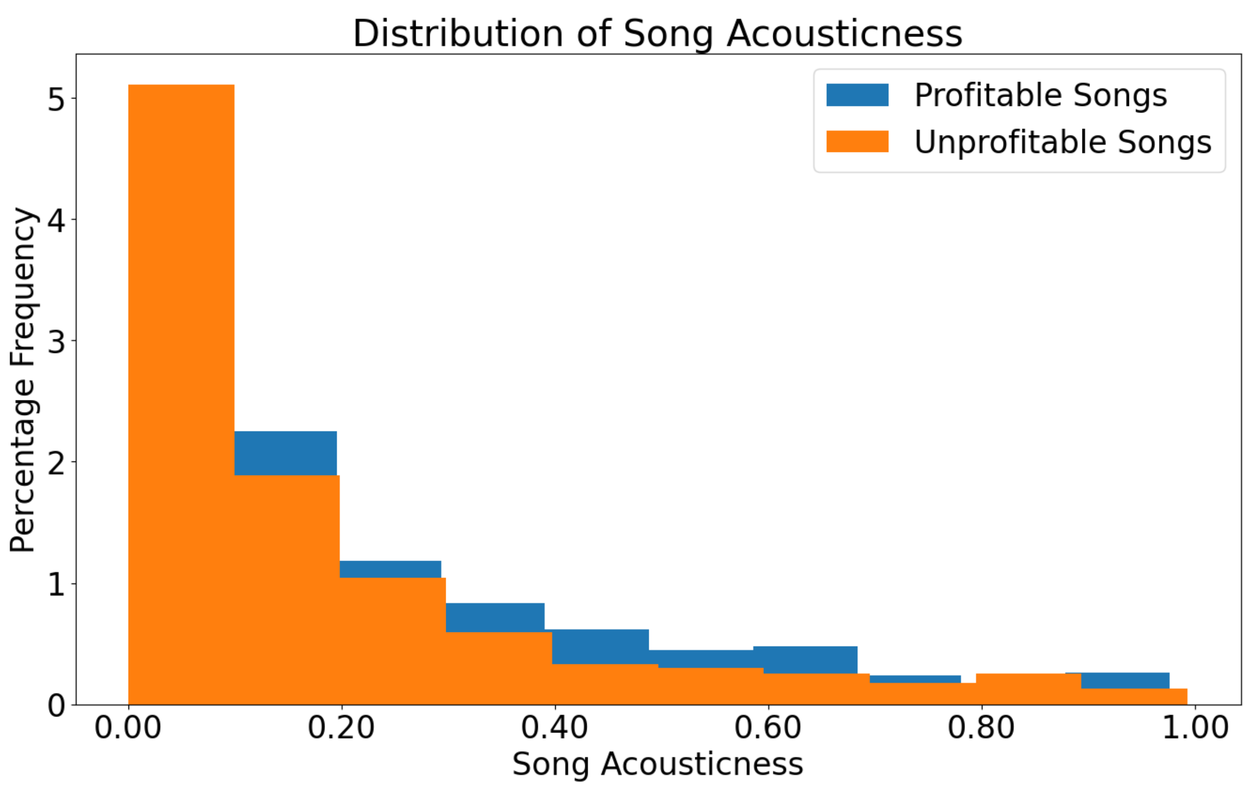

Figure 25: Counterfactual Acousticness - Random Recommendations

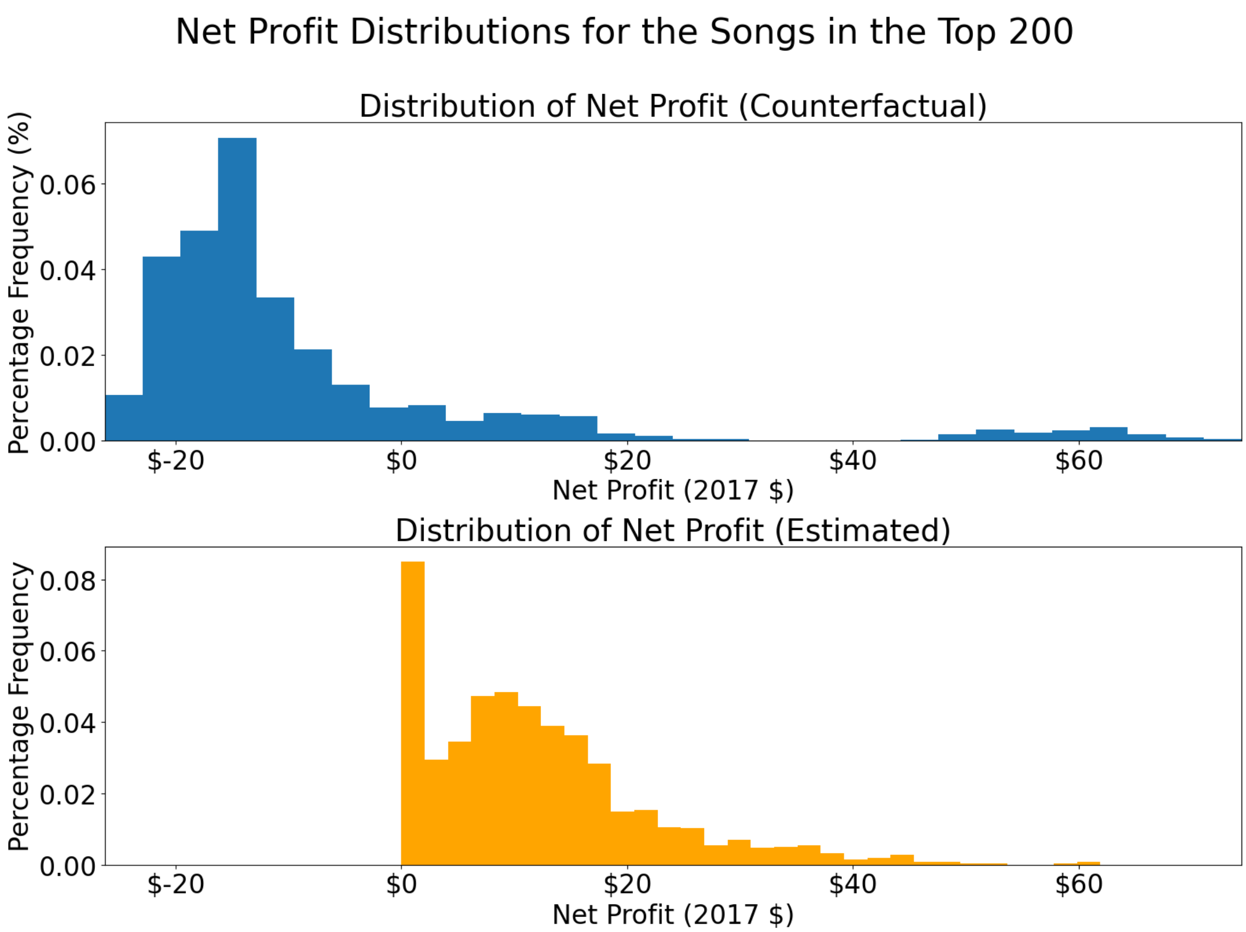

Figure 26: Counterfactual Expected Profit - Popular Recommendations

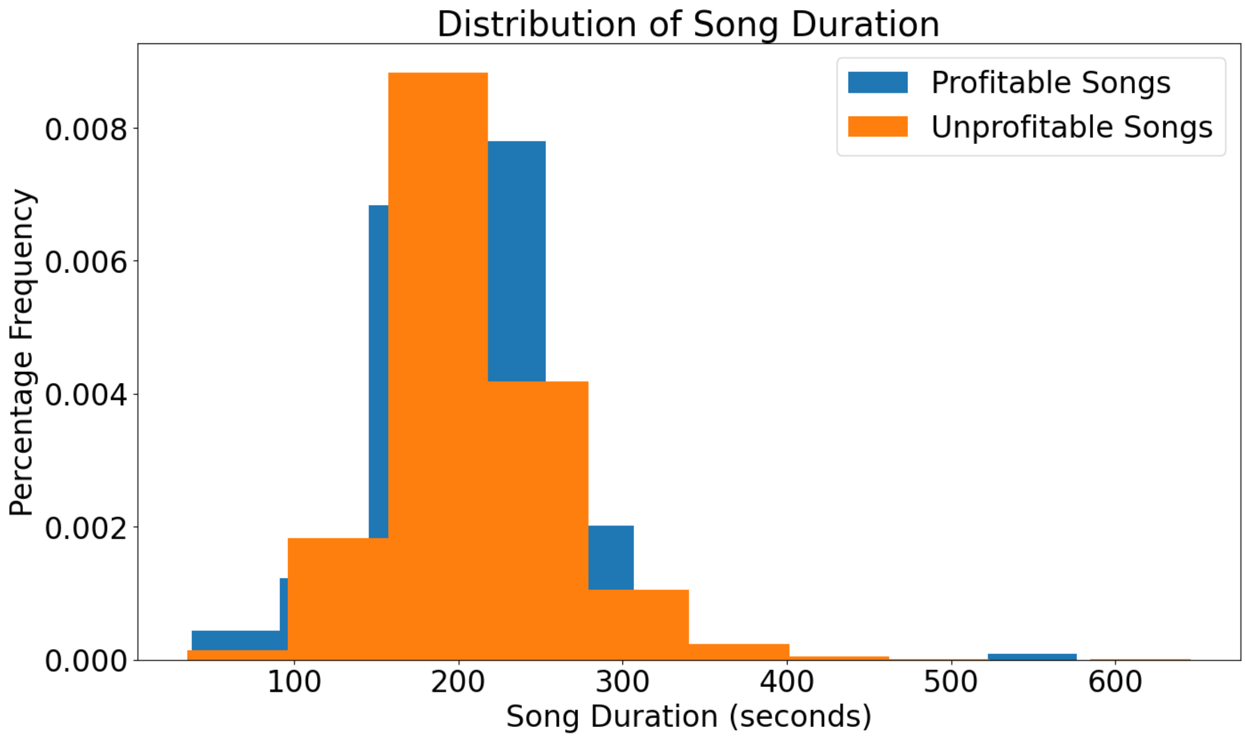

Figure 27: Counterfactual Duration - Popular Recommendations

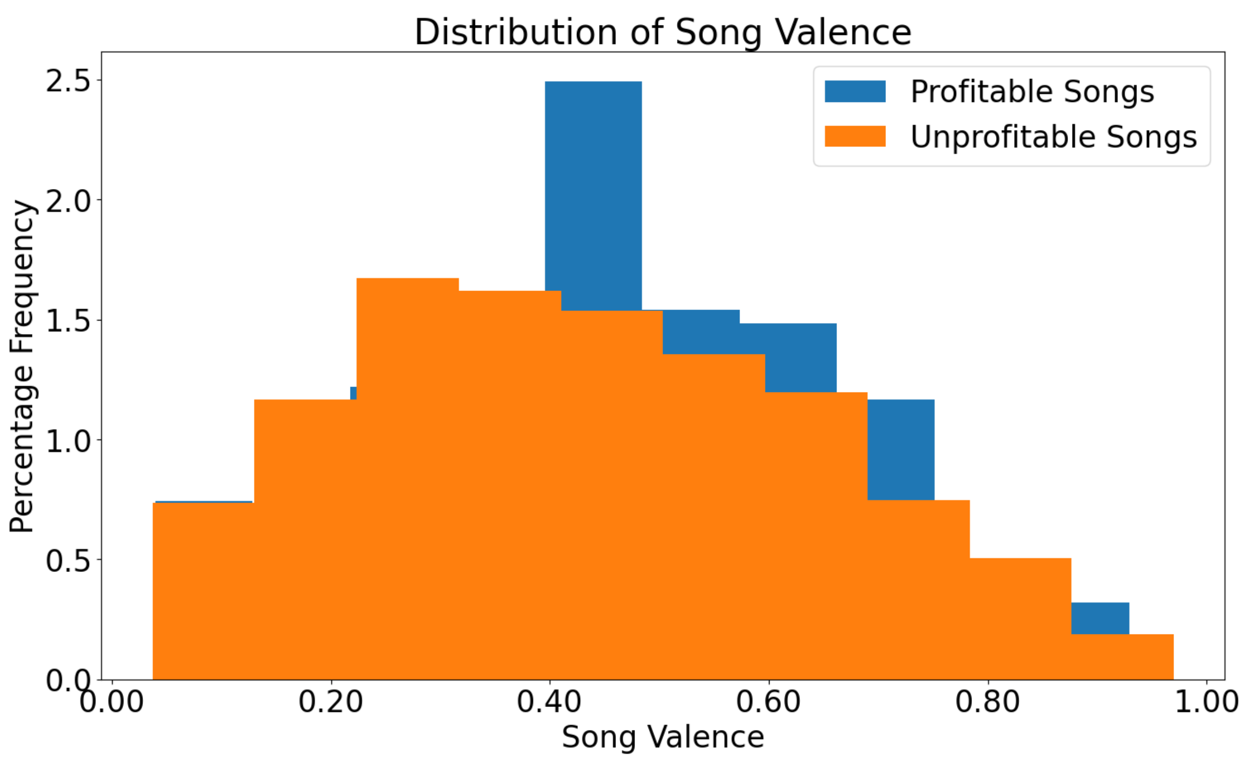

Figure 28: Counterfactual Valence - Popular Recommendations

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.

Calvano, Emilio, Giacomo Calzolari, Vincenzo Denicolò, and Sergio Pastorello. 2020. "Artificial Intelligence, Algorithmic Pricing, and Collusion." *American Economic Review* 110, no. 10 (October): 3267–3297. ISSN: 0002-8282, accessed September 2, 2022. https://doi.org/10.1257/aer.20190623. https://www.aeaweb.org/articles?id=10.1257/aer.20190623.

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.