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

Max E. Schnidman†

May 15, 2026; latest version available here

Abstract

I examine how recommender systems have influenced the music industry and shaped music production. Using a structural model of the recorded music industry, I analyze consumer behavior, platform recommendations, and rightsholder release decisions. I estimate fixed cost of $1.9 million for songs that enter Spotify's Top 200. Counterfactual analysis shows that with randomized recommendations, fewer songs would enter the market, reducing consumer welfare by approximately 2%. The songs that do enter would be 8 seconds longer on average. Popularity-based recommendations that do not account for individual taste would generate a superstar effect—increasing gross profit margins for songs that enter the market to 28%—but reducing consumer welfare by approximately 5%. Although recommender systems have reduced overall variety in music, they have also enabled additional entry and increased consumer welfare.

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

JEL Codes: D43, L15, Z11

*I extend my greatest thanks to my advising committee, Federico Ciliberto, Simon Anderson, and Julie Holland Mortimer. I also extend my thanks to Anton Korinek, Alex Mackay, Jimbo Brand, 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.

†Microsoft, max.schnidman@gmail.com

1 Introduction

Recommender systems, which are designed to match consumers with products they will like, are now the curators of our digital lives, influencing everything from the products we buy to the music we listen to. Such systems are prevalent in many online marketplaces, including Amazon, TikTok, and Netflix, and they have become a key feature of the digital music industry.[1] Music streaming platforms, on which consumers can access a vast catalog of music for a fixed monthly fee (or ads), have become the primary way consumers access music, with streaming accounting for 84% of the recorded music industry's \$17.7bn revenue in 2024.[2] These platforms use recommender systems to generate playlists that surface music to users, and they are where users discover the majority of new music. I investigate how these recommender systems affect the music industry, and how they have shaped the sound of music since their introduction.

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

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

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

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

To estimate this model, I use two primary sources of data: the Music Streaming Session Dataset (MSSD, Brost, Mehrotra, and Jehan 2018), and data from Spotify Charts. The MSSD contains 160 million consumer-level streaming sessions from July to September 2018, including song characteristics, consumer characteristics, listen duration (binned), and whether the song was delivered by a recommender system or through other contexts. Because track identifiers in the MSSD are anonymized, I match them to real Spotify track IDs using audio feature vectors from the Spotify API, recovering the identity of approximately 130,000 songs that also appear in Spotify's Charts catalog. Spotify Charts reports the daily top 200 songs on Spotify for each country, along with stream counts and real Spotify track IDs, providing market-level consumption data for the supply-side estimation. I use the MSSD to estimate consumer demand and recommender system preferences, and the Charts data to estimate the Markov processes governing rival song evolution, compute expected revenues, and estimate fixed costs.

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

Using my structural model, I estimate a gap between consumer demand and recommender systems, driven by differences in their preferences, and find that producers respond to this gap by targeting the recommenders' objectives jointly with consumer preferences. For example, although consumers are likelier to listen to longer songs, the recommender system is likelier to surface shorter songs and producers respond by releasing shorter songs. I also estimate a median fixed cost of $1.9 million for releasing a song into the Top 200 on Spotify.

My counterfactual analysis focuses on changing Spotify's recommender system to see how it has affected song characteristics. I first impose random recommendations, as a proxy for no recommendations, or a naive consumer search process. I find that in the absence of recommender systems, songs are on average 8 seconds longer and less profitable. As a result, fewer songs would be released, and consumer welfare is approximately 2% lower than in the status quo. I also consider a recommender system based exclusively on song popularity, as a proxy for a ban on personalized recommendations, or a simulation of the digital music storefronts of the 2000s (e.g., iTunes). Such a recommender system would generate a superstar effect, which increases gross profit margins for songs that enter the market to 28%, but reduces consumer welfare by approximately 5%. In this counterfactual, songs are 3 seconds shorter on average, and higher-tempo songs are likely to be released. This suggests that Spotify's recommender system has indeed changed the sound of music, and that while these changes have reduced the variety of music available to consumers, they have also increased both the quantity of songs and consumer welfare.

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

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

1.1 Literature Review

I contribute to multiple strands of the economics literature. First, I contribute to research on the economics of music by developing a structural model of the music streaming industry.

Aguiar, Waldfogel, and Waldfogel (2021) use reduced-form analysis to identify bias in the rankings of songs on Spotify's New Music Friday playlist. They find that higher-ranked songs tend to perform better after placement on the playlist, which suggests that curators are looking to maximize streams for their playlist. They also find that the curators of this playlist tend to favor songs by women and from independent labels, because they rank higher than their post-placement performance would suggest. Benner and Waldfogel (2016) use a difference-in-differences design to estimate how the digitization of recorded music has affected the release strategy of record labels. They find that, after digitization, major labels both release fewer albums and become more reliant on previously successful artists; conversely, independent labels release more albums. Aguiar and Waldfogel (2021) estimate the effect of including a song on a Spotify playlist using a regression discontinuity and instrumental variable design.

This paper also builds on Aguiar and Waldfogel (2018), who develop a structural model of the digital music industry. They model consumer demand for digital music across countries and estimate the fixed cost of entry under three scenarios: perfect quality foresight, no quality foresight, and imperfect quality foresight, in which firms know their songs' quality with some forecasting error. They estimate this fixed cost as the expected revenue of the worst-performing song and find that the fixed cost is higher when rightsholders have no quality foresight. Their counterfactual analysis finds that tripling the number of songs available to consumers under imperfect foresight adds nearly 20 times as much consumer surplus as doing so under perfect foresight. I extend this model to the music streaming industry by modifying the choice structure to reflect the streaming industry, incorporating a recommender system in the model, and introducing forward-looking rightsholders. I adjust the fixed cost model to estimate costs from a lognormal distribution, rather than assuming that the fixed cost is the expected revenue of the worst-performing song.

This paper also incorporates the effects of changing business models on labels and artists. Most services use a pro rata model (paying artists by share of total streams), which some argue unfairly favors superstar heavy rotation. Bergantiños and Moreno-Ternero (2025) provides game-theoretic foundations for both the pro rata system and an alternative "user-centric" payout, where each user's fee is split only among the artists they personally listen to. They explore hybrid models that lie between these two extremes, aiming to better align incentives of platforms, fans, and musicians. I explore a similar set of business models, focusing instead on the per-subscriber fee and the per-play fee, and how these models affect song characteristics. This line of research speaks to emerging industry experiments with user-centric payments and other schemes to make streaming income more equitable for diverse artists. Mortimer, Nosko, and Sorensen (2012) examine the impact of file-sharing on sales of recorded music and on the demand for live concert performances. They suggest that while file-sharing reduced album sales, it simultaneously increased demand for concerts, especially for smaller artists.

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

The recommender system literature is divided on the impact of recommender systems on consumer behavior. Fleder and Hosanagar (2009) demonstrate that popular recommendation engines (like collaborative filters) often exhibit a popularity bias that can create a "rich get richer" feedback loop. Their models and simulations show that while individuals may discover new products, recommender systems tend to steer users toward the same set of popular items, which reduces aggregate sales diversity. In other words, personalization can increase a single person's variety but still amplify blockbusters overall. The authors highlight that small design tweaks could mitigate this effect, suggesting platform designers can choose to promote diversity or concentrate attention depending on their objectives. This has implications for market efficiency, as a biased recommender might leave niche products (that some consumers would love) under-recommended, resulting in missed matches. I test their hypothesis in the context of the music industry, where recommender systems are now ubiquitous, using counterfactual analysis.

On the other hand, some empirical studies find that recommender systems can broaden user horizons. For example, an analysis of iTunes social-network data by Hosanagar et al. (2014) found that personalized music recommendations actually widened consumers' exposure to new artists and fostered more overlap in what people listened to (counter to "filter bubble" fears). Users who received tailored song suggestions shared more common interests and discovered music outside their initial comfort zone. This suggests recommender systems need not fragment audiences into isolated niches; with well-calibrated algorithms, they can increase total consumer welfare by helping listeners find content they enjoy but would not have found on their own. The net impact likely depends on the platform's goals (e.g., maximizing click-through vs. encouraging exploration) and the specific design of the recommendation algorithm.

From an economics perspective, recommender systems can also affect competition among firms and raise policy questions. Fletcher, Ormosi, and Savani (2023) identify several systemic biases in recommendation algorithms – including popularity and incumbency bias (favoring well-known items), homogeneity bias (recommending very similar content), and conformity bias (herding users toward mainstream choices). They argue these biases may collectively harm market efficiency by skewing consumer choices and making it harder for new or diverse providers to compete. Biased recommendations can increase concentration, create higher entry barriers for new products, and reduce variety, even if the platform is not intentionally anti-competitive. Such findings have led to calls for more transparent and accountable algorithms. Proposed remedies include incorporating diversity or fairness objectives into recommendation criteria, or regulatory oversight to ensure platforms aren't locking in consumers unfairly. This intersection of algorithm design and competition policy is a nascent but important area, recognizing that recommender systems now serve as key intermediaries in digital markets.

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

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.

The recorded music industry has undergone a dramatic transformation over the past three decades. The rise of the internet in the 1990s facilitated widespread digital piracy through services like Napster, leading to a significant decline in industry revenues. In response, Apple launched the iTunes store in 2003, establishing a legal market for digital

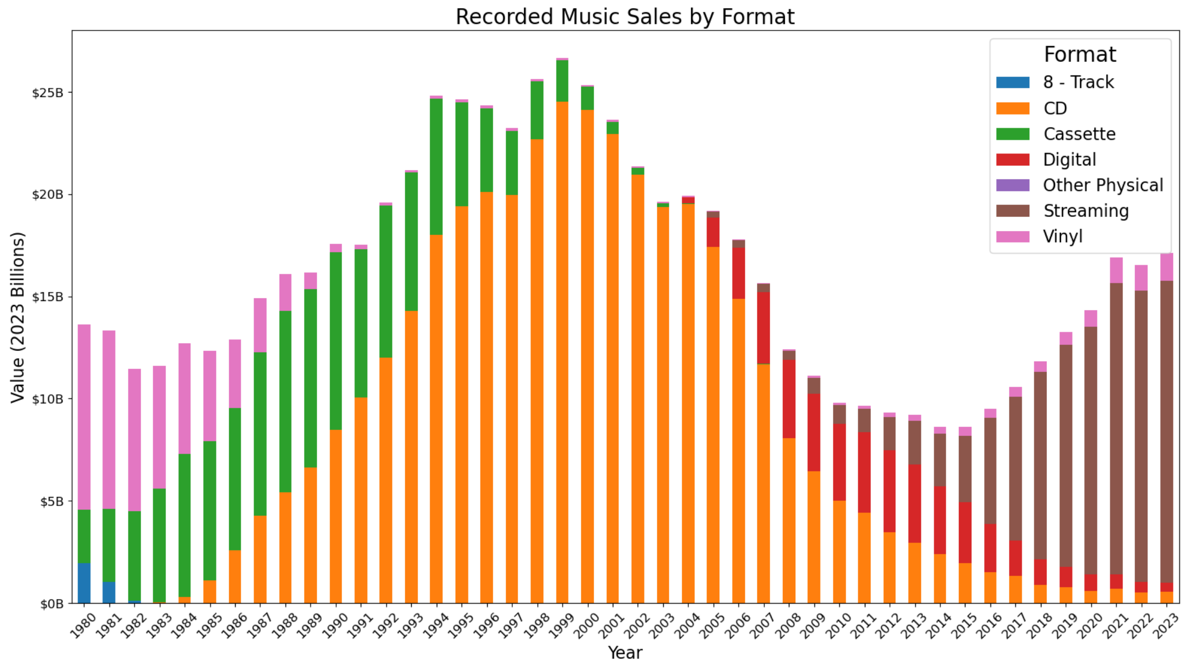

Figure 1: Real Revenue of the Recorded Music Industry, 1990-2023

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

music downloads and pioneering the unbundling of albums into individual tracks. The pivotal shift occurred in the early 2010s with the mainstream adoption of streaming services, led by Spotify. These platforms offered consumers on-demand access to vast music catalogs for a subscription fee, effectively combating piracy and reversing the industry's financial decline. Today, streaming has become the dominant mode of music consumption, accounting for 84% of industry revenue in 2023 (Figure 1), and fundamentally reshaping how music is produced, discovered, and monetized.

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

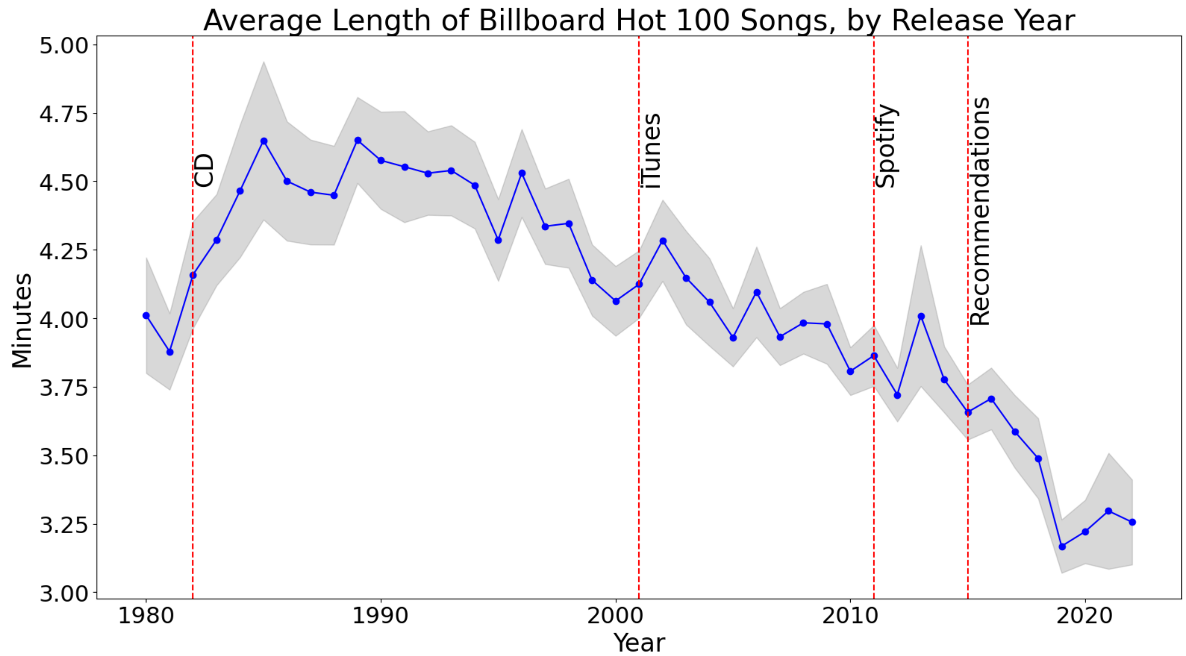

Figure 2: Average Song Duration on Billboard's Hot 100, 1990-2022

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

(e.g., danceability, energy, valence) in my model. I include descriptions of these characteristics, and examples in popular songs, in the Appendix (see Tables 19 and 20).

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.[6] 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. In Appendix 8, I estimate the correlation between the introduction of new music formats and song duration, finding that the introduction of streaming services and recommender systems correlates with a 40-second decrease in the average length of songs on Billboard's Hot 100. In Appendix 8, I also examine whether consumer preferences have changed over time, and whether these preferences are driving the changes in song length, finding that consumer preferences for song length have not significantly changed over time.

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

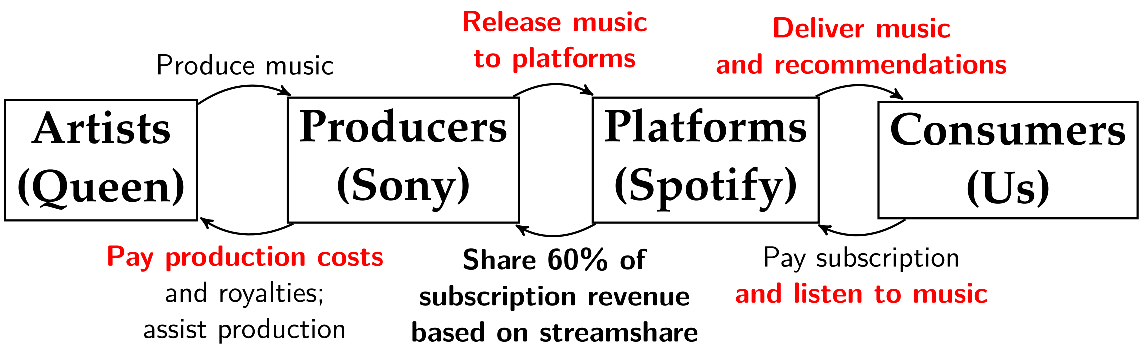

Figure 3: Vertical Structure in the Music Industry

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

2.2 Industry Structure

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

Beginning on the left, artists create music, either by themselves or in contract with a record label, who serves as a rightsholder. An artist on contract with a rightsholder typically receives an advance and production assistance in exchange for ownership over the music they create. Artists also receive a share of the revenue (royalties) from the music they create, as negotiated with the rightsholders.[7] 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.[8]

Rightsholders, such as Sony, Warner, and Universal (the Big Three record labels), are responsible for distributing music to consumers, either through physical media (e.g., CDs) or through digital platforms (e.g., Spotify). They also search for new and upcoming artists to sign to contracts and promote their music. These labels also have a wide variety of subsidiary labels (or sub-labels) to focus on particular types of music or audiences. These sub-labels 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

7. 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

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

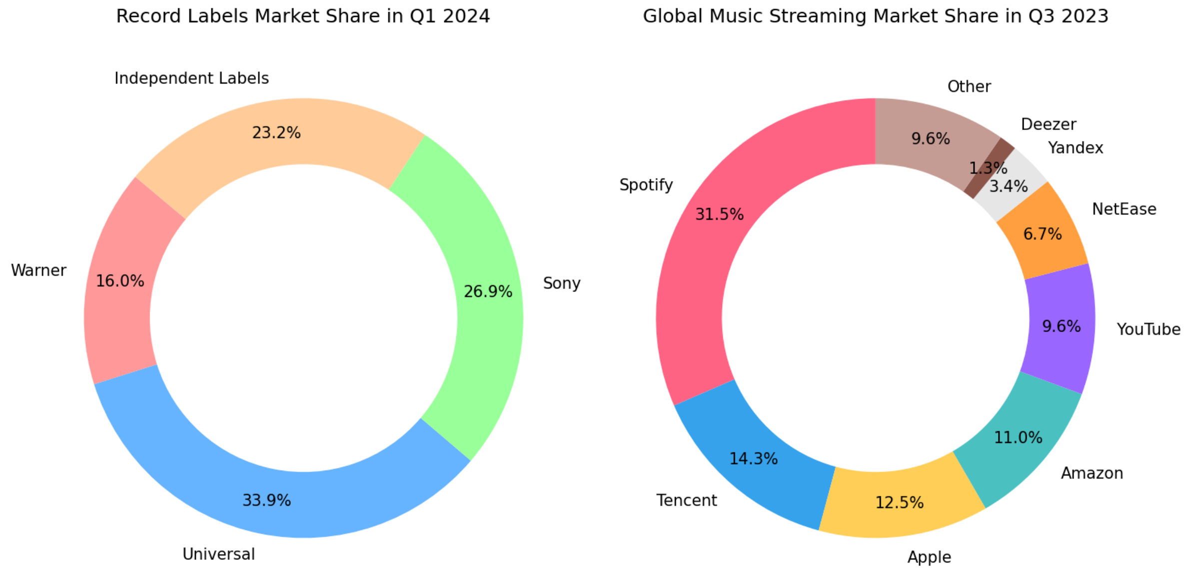

Figure 4: Concentration in the Recorded Music Industry, 2023

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

and streaming platforms in more detail in the following subsection.

Rightsholders are a highly concentrated section of the industry, with the Big Three (WMG, Sony, and UMG, including their sublabels) capturing 77% of the market. Other independent labels comprise the remaining 23% of the market. Figure 4 shows the market share of rightsholders (and streaming services).

Streaming platforms, such as Spotify, Apple Music, and Amazon Music, are 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. They have revolutionized the recorded music industry, allowing consumers to access a vast catalog of music for a fixed monthly fee. As with rightsholders, this section of the industry is highly concentrated, with five firms comprising approximately 80% of the market. Figure 4 shows the market share of streaming platforms (and rightsholders).

These platforms have relatively homogeneous music catalogs, hosting songs from the Big Three and many independent labels. Instead, they differentiate instead on their recommendation engines, interface, and ancillary features (e.g., exclusive podcasts, integration with smart devices, etc.). I speculate that the presence of YouTube as a free, ad-supported platform for music and lyric videos made it difficult for these platforms to compete on exclusive content.[9] 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 options to consumers, which I group into two types: ad-supported access and premium subscriptions. Ad-supported access 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.[10] Premium subscribers can also download songs for offline listening, stream higher quality audio, and listen to 15 hours of audiobooks per month.

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.[11] Spotify pays rightsholders for royalty-bearing streams (RBS), defined as any play of a song that lasts more than 30 seconds.[12] 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

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

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

11. Singleton (2015)

12. 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

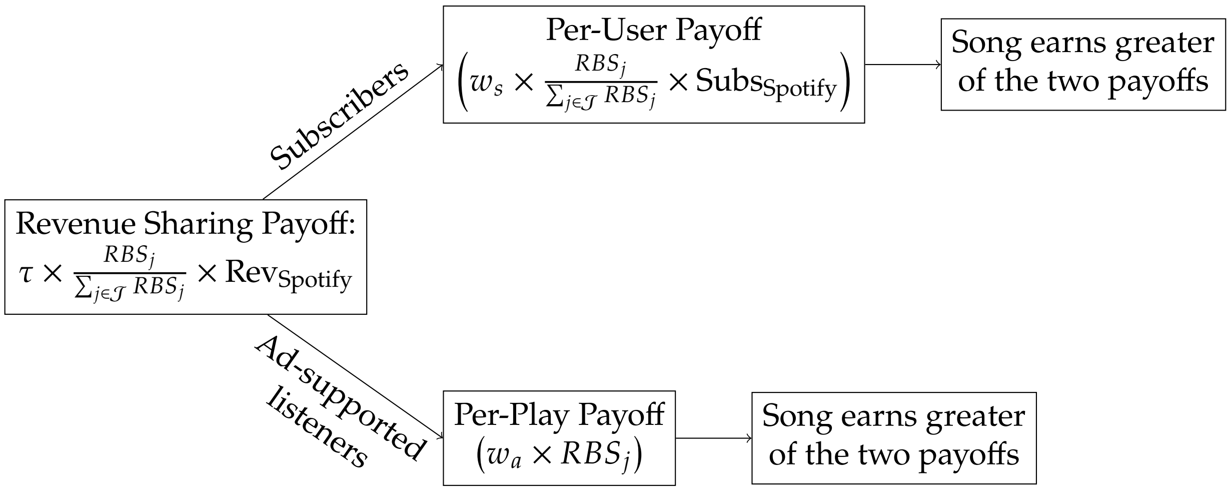

Figure 5: Revenue Sharing Payoff Structure

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

pays rightsholders the greatest of a share of ad revenue of a per-stream fee. Figure 5 shows the payoff structure for rightsholders.

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

At launch, Spotify charged $9.99 for a premium subscription, so the revenue share and per-subscriber fee were equivalent at that time. Since Spotify has gone public in 2017, its premium average revenue per user has been well below the per-subscriber fee, primarily because of family and student plans, which reduce the price per user. Assuming that Spotify has not renegotiated the per-subscriber fee with rightsholders, this would imply that this fee (times the number of subscribers) is greater than the revenue share, and that Spotify is paying rightsholders the per-subscriber fee. Singer and Rosenblatt (2023) suggest, however, that the per-subscriber fee is a floor, and that Spotify pays rightsholders a revenue share of approximately 65% of gross revenue.[13]

The structure of this contract is vital for understanding the incentives of rightsholders to release different kinds of music on Spotify. Firms have a clear incentive to reduce song length to increase the number of RBS, and thereby increase their streamshare and

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

revenue from Spotify. Spotify, however, would pay more for ad-supported subscribers if more streams occurred, so they would prefer to have longer songs. Consumers also have preferences over song length, which can affect these incentives.

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

3 Data

I leverage the Music Streaming Sessions Dataset (MSSD, Brost, Mehrotra, and Jehan 2018), and augment it with data from Spotify Charts.

3.1 Music Streaming Sessions Dataset

The MSSD consists of approximately 160 million consumer-level streaming sessions between July 15th and September 18th of 2018, containing roughly 3.7 million unique songs and 2 billion song-consumer interactions. The MSSD defines a streaming session as any listening session with less than 60 seconds between songs. The data contain only streaming sessions with at least ten songs, and truncate all streaming sessions after twenty songs.

The MSSD contains song-level characteristics for each of the approximately 3.7 million songs in its catalog. These 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 to count as an RBS. I also observe details about the consumer's streaming session: the position of the song in the session, the date and hour when they listened to each song in the session, and whether the consumer was listening to a song they searched for, their own collection, an editorial playlist, or an algorithmic playlist or radio station. Additionally, I observe what the consumer did after each song, which I use to determine under what circumstances a consumer ended their streaming session. Moreover, I observe the consumer's subscription status at the time of listening. I use these choice-level data to estimate my model of consumer demand and the recommender system.

3.1.1 Track Unmasking

A key challenge with the MSSD is that all track identifiers are anonymized: each song is labeled with a randomly generated UUID rather than its real Spotify track ID. This anonymization prevents direct linkage to external metadata, chart performance, or the Spotify API. However, the MSSD does provide a vector of audio features (danceability, energy, loudness, tempo, etc.) for each anonymized track.

I exploit this feature information to recover the identity of songs through a matching procedure. My second data source, Spotify Charts, reports the top 200 songs on Spotify daily for each country, including each song's real Spotify track ID. Using the Spotify API, I retrieve the same set of audio features for each charting song, creating a reference dataset of known songs with both real identifiers and feature vectors.

I then match anonymized MSSD tracks to their real identities in two phases. First, I perform exact matching on feature vectors: tracks whose features are identical across all dimensions are matched deterministically. Second, for remaining unmatched tracks, I compute Euclidean distances between the normalized feature vectors of anonymized and known tracks, accepting matches only when the distance falls below a strict threshold (10^{-6}). This procedure recovers the real identity of approximately 130,000 songs in the MSSD that also appear in the Spotify Charts catalog.

3.1.2 Session Sampling

The full MSSD contains approximately 160 million sessions across 660 daily CSV files, which is too large for estimation. I draw a representative sample using a proportional stratified sampling strategy. First, I identify all sessions that contain at least one unmasked song, ensuring that each sampled session can be linked to externally verifiable song characteristics. I then allocate a target of 100,000 sessions proportionally across the 66 days in the data, so that each day is represented in proportion to its number of eligible sessions. Within each day, I randomly sample the allocated number of sessions and retain all rows (songs) belonging to those sessions, preserving complete streaming histories.

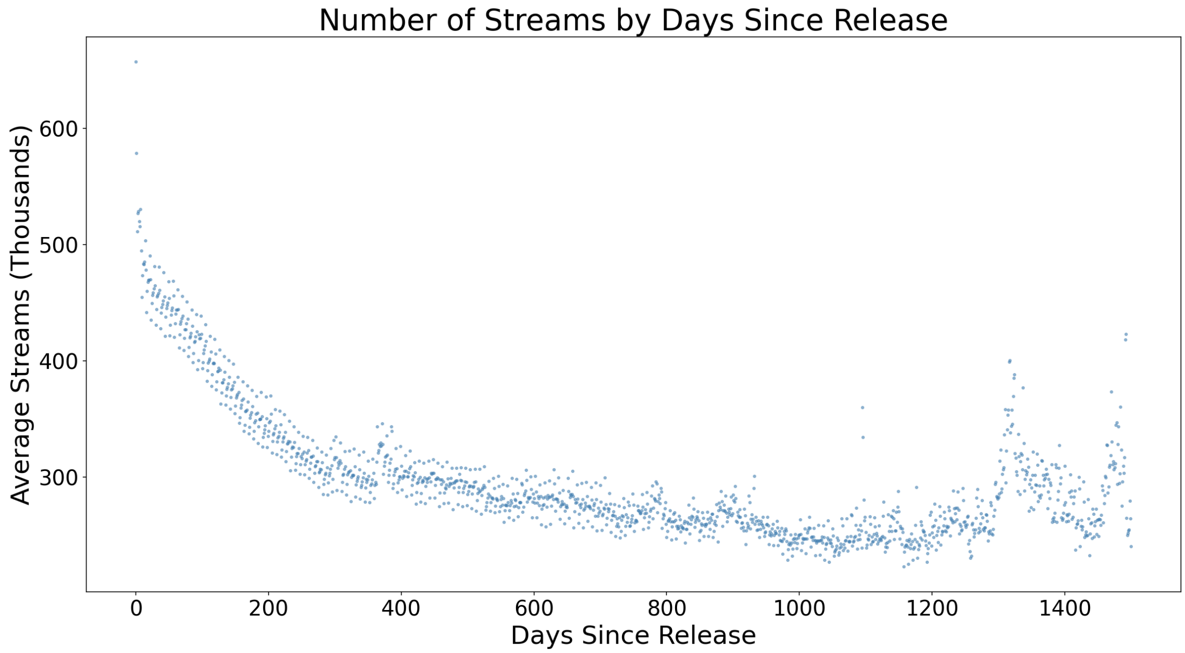

Figure 6: Number of Streams of Songs on Spotify's Top 200, by Days since Release

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

This procedure yields a sample of 100,000 sessions containing approximately 130,000 unique songs and 2 million song-consumer interactions. I use this sample for all demand and recommender system estimation.

My second data source is Spotify Charts, which reports the daily top 200 songs on Spotify for each country, including stream counts and real Spotify track IDs.[14] I use the Charts data to construct the Markov processes for rival song characteristics, compute expected revenues, and estimate fixed costs in the supply model.

3.2 Descriptive Statistics

The Spotify Charts data contains $9{,}245$ unique songs that entered the US top 200 between 2017 and 2021; descriptive statistics for these songs are reported in Appendix Table 18. The Charts data also provides information about the lifecycle of songs. Figure 6 reports the number of streams of a song by day after release:

This figure shows the average number of streams each song that made it on Spotify's Top 200 received in the days since its release. Unsurprisingly, songs get a significant num-

14. I rely on a Kaggle dataset that scraped Spotify Charts and the Spotify API: https://www.kaggle.com/edumucelli/spotifys-worldwide-daily-song-ranking

	Mean	Median	Standard Deviation	Min	Max
Duration (s)	219.31	213.00	68.07	30.01	1753.77
Release Year	2012	2016	10	1950	2018
Acousticness	0.25	0.13	0.27	0.00	1.00
Danceability	0.62	0.64	0.16	0.00	0.99
Energy	0.64	0.66	0.21	0.00	1.00
Instrumentalness	0.07	0.00	0.21	0.00	1.00
Liveness	0.20	0.13	0.17	0.00	1.00
Loudness	-7.47	-6.71	3.79	-60.00	2.68
Mode	0.62	1.00	0.49	0.00	1.00
Speechiness	0.13	0.07	0.14	0.00	0.97
Tempo (BPM)	121.27	120.09	29.49	0.00	235.86
Time Signature	0.97	1.00	0.16	0.00	1.00
Valence	0.48	0.47	0.24	0.00	1.00

Table 1: MSSD Song Characteristics ($N = 129,549$ songs)

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

ber of their streams in the first 100 days after release, with the average number of streams above $400,000$ per day for the first 100 days. After that, the number of streams decreases, with a small uptick around the one and two-year marks, but continuing to fall off over time. The number of streams becomes more volatile after the three-year mark, because fewer songs have been out for that long in my data. Table 1 reports the descriptive statistics for the songs in the Music Streaming Sessions Dataset.

My sample of the MSSD contains approximately 130,000 unique songs, with an average length of 3 minutes and 39 seconds, with a standard deviation of 1 minute and 8 seconds. Compared to the Spotify Charts data, these songs are longer and have a higher standard deviation in length. These songs are also older than the Spotify Charts songs, with an average release year of 2012 (median 2016), compared to 2017 (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 sample 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 sample variables. Table 2 reports the consumer-level statistics for my sample of the Music Streaming Sessions Dataset.

Consumers in my sample are primarily premium subscribers, with 79% of the sample

	Mean	Standard Deviation
Session Length	17.67	3.37
% Premium Subscribers	0.79	0.41
% RBS	0.57	0.50
% Completion	0.33	0.47
% Morning Listen	0.23	0.42
% Afternoon Listen	0.39	0.49
% Evening Listen	0.30	0.46
% Night Listen	0.08	0.27
% Monday Listen	0.15	0.36
% Tuesday Listen	0.15	0.36
% Wednesday Listen	0.14	0.34
% Thursday Listen	0.14	0.34
% Friday Listen	0.15	0.36
% Saturday Listen	0.13	0.34
% Sunday Listen	0.13	0.34
% Catalog Listen	0.18	0.39
% Chart Listen	0.02	0.15
% Editorial Playlist Listen	0.19	0.39
% Algorithmic Playlist Listen	0.01	0.12
% Algorithmic Radio Listen	0.13	0.34
% User Collection Listen	0.46	0.50

Table 2: MSSD Consumer Characteristics ($N = 2m$ song-consumer interactions)

Consumers in the MSSD have long streaming sessions. 57% of interactions qualify as a Royalty-Bearing Stream (30+ seconds), but only 33% result in full song completion. Consumers primarily listen to their own collections, with about 14% of listens algorithmically driven.

being premium subscribers. This is significantly higher than the percentage of premium subscribers Spotify reports, which is 40% of its user base.[15] It is, however, more representative of the percentage of revenue Spotify earns from premium subscribers, which is 88% of its revenue.[16] These users have very active streaming sessions, with an average session length of 18 songs. Of all consumer-song interactions, 57% last long enough to qualify as a Royalty-Bearing Stream (at least 30 seconds), while only 33% result in full song completion. The distinction matters because the demand model uses the RBS threshold as its outcome variable, while the recommender system model uses song completion. Listening time is even throughout the week, with 13-15% of sessions occurring on each day of the week. Within a day, however, very little listening occurs at night (12-6 AM), with only 8%

15. Spotify Q2 2024 Earnings Report
16. Spotify Q2 2024 Earnings Report

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 64% of sessions being from these sources. Algorithmic playlists and radio stations consist of 14% of streaming sessions. Editorial (human-curated) playlists and top charts account for 21% of sessions. I use these separate listening contexts to separately estimate the preferences of consumers who receive songs through the recommender system (algorithmic playlists or radio) and those who do not.

4 Model

To evaluate the effect of recommender systems on the music industry, I develop a structural model of the industry, with three sets of agents: consumers, a recommender system, and rightsholders. Consumers (the demand side) receive songs from the platform (and its recommender system) and choose whether to listen to them. I capture this choice using a random utility model, which generates a probability of listening to a song based on its characteristics and the consumer's characteristics. The recommender system, which I treat as an exogenous technology, computes the probability consumers receive particular songs based on their characteristics and the consumer's characteristics, and it surfaces songs in proportion to their probability of being listened. The joint probability of being surfaced and the probability of being heard, times the number of potential listeners, is the demand rightsholders face. On the supply side, rightsholders decide whether to release songs provided to them by artists, paying a fixed cost to releasing them. They enter the market if the expected revenue, which is a function of the choice probabilities at the time of release and in the future, from the song is greater than the fixed cost of releasing it. These rightsholders are forward-looking, anticipating the evolution of the market and the recommender system through first-order Markov processes. They draw their fixed cost from a lognormal distribution, whose parameters are functions of song characteristics and known to the rightsholder. My solution concept is an oblivious equilibrium, where each firm considers only the long-run average choice of the industry, rather than each rival's choice. Figure 7 describes the timing of the model each period.

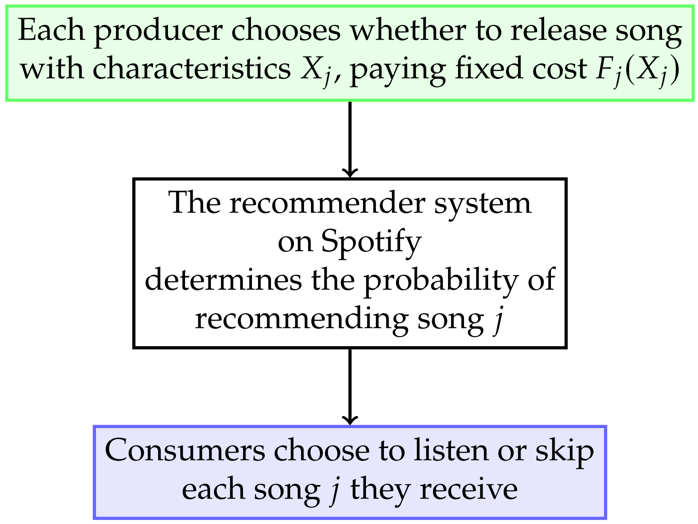

Figure 7: Timing of the Model in Each Period

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

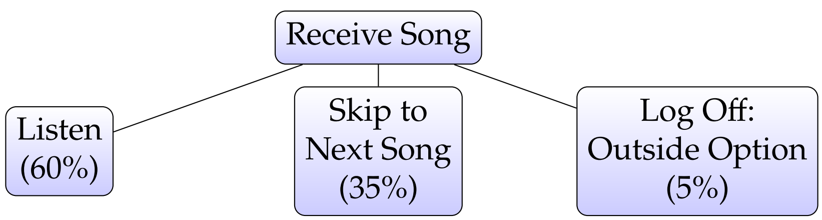

Figure 8: Consumer Decision Tree

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

4.1 Demand

Consumers in my demand model are subscribers to a streaming platform offering them a catalog of songs.[17] Each day, these consumers open the streaming app and start receiving songs from the platform, whether from the recommender system or from other contexts. For each song they receive, consumers make one of three possible choices: listen to the song up to the amount necessary for a Royalty-Bearing Stream (RBS), skip the song, or stop listening to the platform, which I treat as an outside option. Figure 8 describes the decision tree for consumers in the demand model.

I maintain one assumption about consumers in my model:

17. 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.

Assumption 1 *Consumers do not consider how their choice affects future personalized recommendations.*[18]

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. Were I to relax this assumption, consumers would face a dynamic problem, incorporating their perception of the recommender system into their decision-making process. Consumers would strategically select songs they believe would improve their recommendations, particularly early in their streaming session, when they are more likely to be uninformed about the recommender system's characteristics. That is, they would listen in full to songs they like, but decline to complete songs they dislike (in terms of the characteristic space). This interacts with the RBS cutoff, as consumer may still listen to a song up to the RBS cutoff, but not to completion. As such, firms may have a strategy of releasing songs that are a poor fit for consumers, in order to get RBSs as consumers train their recommendations. The tradeoff, however, is that these songs would receive far less revenue in the long term.

Alternatively, I could relax this assumption by allowing consumers to choose a bundle of songs to listen to, rather than individual songs. In practical terms, they would be selecting a playlist of songs, or a radio station. Here, they would consider the utility of the entire bundle, and how the bundle in aggregate would affect their recommendations. The modeling challenge would be capturing complementarities of the bundle, and the RBS decisions for each song in the bundle. I capture some bundle selection effects indirectly through consumers' adaptive expectations of songs, as I discuss in section 4.2.

4.1.1 Choice Structure and Utility Model

Formally, consumers i have access to a set of songs $\mathcal{J}$ that they receive from the platform. Each song $j \in \mathcal{J}$ has a vector of characteristics X_j, which are known to the consumer. Consumers have preferences over these characteristics. Each consumer, when receiving one of these songs from the platform, faces three alternatives B. First, they can choose to listen to the song, which I denote as L, receiving the utility from listening to the song. The song's characteristics enter into the consumer's utility from listening, as does the consumer's own characteristics Y_i. Second, they can skip the song, which I denote as S, receiving the utility from skipping the song. I discuss the utility of skipping songs in section 4.2. Finally, they can log off the platform, which I denote as O, earning zero utility.

In my model, consumers choose alternative B if and only if

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

$$U(B, Y_i, X_j) \geq U(B', Y_i, X_j) \text{ for all } B' \in \mathcal{B} \setminus B$$

In addition to the above deterministic utility, I incorporate a random error term ϵijB into the utility function, to capture the unobserved heterogeneity in the utility of each alternative. To formally model this discrete choice behavior influenced by both observable characteristics and unobserved heterogeneity, I employ a random utility model.

I now describe the functional form of the utility generated by each alternative B. Consumer i's utility of listening to a particular song j in session position s is given by:

$$U_{L,ijs} = \beta_i X_j + \gamma_L Y_i + \eta_L \log(s) + \epsilon_{ijs} \tag{1}$$

In this utility function, X_j are a vector of linear and quadratic song characteristics (variables specific to each song j, or alternative), Y_i are a vector of consumer characteristics (variables specific to each consumer i across all choices B, or case), $\eta_L \log(s)$ captures position effects through the log of session position, and ϵ_{ijs} is a Type 1 (Gumbel) Extreme Value error term. The consumer characteristics and session fixed effects are case-specific variables, so their parameters are also case-specific, following the standard in conditional choice models (Train 2009). Intuitively, consumers prefer certain types of music, which I decompose into quantitative characteristics, and their utility from a particular song may depend on when they are listening, both during the day, and where they are in their streaming session. Additionally, consumers are likely to have heterogeneous preferences over song characteristics, which I capture by allowing the parameters on song characteristics to vary across consumers. Moreover, 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, searching for 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:

$$U_{i0s} = \epsilon_{i0s} \tag{2}$$

4.2 Utility of Skipping Songs

To capture the utility of skipping to the next song, consumers form adaptive 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:

$$U_{S,ijs} = \beta_i E_{is}[X_j | X_{j,s-1}] + \gamma_S Y_i + \eta_S \log(s) + \epsilon_{ijs} \tag{3}$$

I refine these expectations using 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.

4.2.1 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 assumption of a Type 1 Extreme Value (T1EV) error term leads directly to the familiar conditional logit probability structure. The probability that consumer i listens to song j in session position s, conditional on the song being recommended, is given by:

$$\begin{aligned} & P(i \text{ listens to } j) \\ = & \frac{\exp(\beta_i X_j + \gamma Y_i + \eta_s)}{1 + (\exp(\beta_i X_j + \gamma Y_i + \eta_s) + \exp(\beta_i E_{is}[X_j | X_{j,s-1}] + \gamma Y_i + \eta_s))} \end{aligned} \tag{4}$$

The probability is the exponentiated utility of listening divided by the sum of exponentiated utilities of all available options (Listen, Skip, and the Outside Option, whose utility is normalized to 0, yielding $exp(0) = 1$ in the denominator).

Realistically, however, consumers will have different preferences based on their listening context. This distinction is vital because consumers actively seeking music (direct selection) may exhibit different preferences or sensitivities to song characteristics compared to when they are passively consuming recommendations. As such, I allow for different preference parameters for consumers who are using the recommender system and those who are not. This allows me to capture the difference between high-information consumers, who are actively seeking music, and low-information consumers, who are passively consuming recommendations. Consumers who are using the recommender system have the following choice probability:

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

Consumers who are not using the recommender system, whom I refer to as direct selection, have the following choice probability:

$$\begin{aligned} &P(i \text{ listens to } j|\text{RS does not surface } j \text{ to } i) \\ &= \frac{\exp(\beta_{d,i}X_j + \gamma_d Y_i + \eta_{d,s})}{1 + (\exp(\beta_{d,i}X_j + \gamma_d Y_i + \eta_{d,s}) + \exp(\beta_{d,i}E_{is}[X_j|X_{j,s-1}] + \gamma_d Y_i + \eta_{d,s}))} \end{aligned} \tag{6}$$

Information on which consumers are conditioning on the recommender system comes from the listening context information in the MSSD, as described in Table 2. Consumers who do not use a recommender system have the same utility functions as described above, but their preference parameters (β, γ, η) will be different.

4.3 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 surfaces 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. [19] 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. In a real-world example, Amazon uses collaborative filtering when recommending products "people like you also bought". 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. Continuing the Amazon example, they use content-based recommendations when describing "similar products". 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.[20] They use a combination of user and song characteristics to recommend songs to users. While the recommender system itself is a closely held black box, various research papers have discussed its mechanisms, and I use these papers for guidance in constructing my model of the recommender system, particularly McInerney et al. (2018).

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

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

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

19. Google Developers
20. "Understanding Recommendations," Spotify Safety and Privacy

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. This is because most consumers likely already have extensive, albeit unobserved, listening histories on Spotify. Additionally, anthropological evidence suggests that Spotify does not take very long to move to an exploitation-heavy strategy. (Eriksson et al. 2019)

I further assume that the recommender system does not meaningfully change its parameters, objective function, or strategy in the timeframe of my data. My data only cover two months, making it unlikely Spotify significantly retrained its model in that interval. If Spotify did retrain its model within the interval of my data, I would likely observe in the distribution of songs recommended to consumers, as well as the distribution of songs listened to by consumers.

I estimate the recommender system using the following equation:

$$P(\text{RS surfaces } j \text{ to } i) = \frac{\exp(\eta_1 X_{1j} + \eta_2 \prod_{n=2}^{8} X_{nj} + \eta_3 Y_i + \eta_4 X_j Y_i + \eta_5 X'_{i,j,t} X_{i,j,t-1})}{1 + \exp(\eta_1 X_{1j} + \eta_2 \prod_{n=2}^{8} X_{nj} + \eta_3 Y_i + \eta_4 X_j Y_i + \eta_5 X'_{i,j,t} X_{i,j,t-1})} \quad (7)$$

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. $X_j Y_i$ are interactions between song and consumer characteristics, and $X'_{i,j,t} X_{i,j,t-1}$ are interactions between the characteristics of the current song and the average characteristics of previously listened songs, to capture some aspects of collaborative filtering. η_1, η_2, η_3, η_4, and η_5 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 7 to the MSSD data.

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

$$P(i \text{ listens to } j) = P(\text{RS surfaces } j \text{ to } i) \times P(i \text{ listens to } j|\text{RS surfaces } j \text{ to } i)$$
$$+ (1 - P(\text{RS surfaces } j \text{ to } i)) \times P(i \text{ listens to } j|\text{RS does not surface } j \text{ to } i) \tag{8}$$

For producers, consumers can access their songs in two ways: through the recommender system, or through direct selection. I treat these other means as the complement to the probability the recommender system surfaces the song. Intuitively, it can also represent a function representing consumer awareness of the song outside the recommender system. As previously discussed, because I observe consumer listening contexts, I can separately estimate these choice probabilities.

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. Additionally, I have separate estimates for aware and unaware consumers of music, rather than a single estimate for consumers with an ad-based awareness function.

4.4 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.[21]

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 rightsholders in my model:

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

21. I treat revenue from Spotify as exogenous, because I do not model Spotify as a strategic agent. I also focus exclusively on the subscription business model.

This assumption allows me to model the rightsholder's decision as a one-time binary choice, rather than a multi-product competition problem. It also simplifies the state space of the model, as I do not need to consider potential cannibalization between songs, or segmented competition. Following other literature on dynamic games, I assume that each potential song entrant is short-lived, with only one possible release window. (Weintraub, Benkard, and Van Roy 2008)

If rightsholders were multiproduct firms, then they may be less likely to release newer songs, because of the risk of cannibalizing their vintage songs. Additionally, if they had endogenous release dates, then they may be more inclined to delay releases to periods when competition is less intense, or to try to coordinate releases with other rightsholders. This would create a more complex model, as rightsholders would need to consider the release dates of their rivals, and the potential for coordination in releases.

4.4.1 Expected Revenue

This decision to release hinges on comparing the expected lifetime revenue against the fixed cost of the song, F_j. Fixed costs are the costs of writing, producing, and releasing a song.

In making this decision, rightsholders are forward-looking agents, considering both current and potential future profits when making this release decision against the fixed cost F_j. Once released, a song remains on the platform in perpetuity[22], potentially generating revenue in future periods based on its streamshare.[23]

To effectively make this decision, they must have some way to forecast future period profits. Specifically, rightsholders need to model two sets of state variables/evolutionary processes:

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

I define $\mathcal{X}_t$ as the mean characteristics of all songs on a given day on Spotify Charts, and I define ϕ as the probability the recommender system recommends a song to a consumer in future periods. These variables define the state space for each rightsholder in

22. In practice, songs can be removed, but for the model's horizon, perpetuity is a reasonable simplification.

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

the market, when solving their dynamic problem. I model the evolution of these two state variables using a first-order Markov process, where the state variables are a function of their previous values and a stochastic error term:

$$\mathcal{X}_{t+1} = \nu_0 + \nu_1 \mathcal{X}_t + \epsilon_t^{\mathcal{X}} \tag{9}$$

$$\phi_{j,t+1} = \psi_0 + \psi_1 \phi_{jt} + \epsilon_{jt}^{\phi} \tag{10}$$

Rightsholders do not have full awareness of the recommender system's potential changes, so I assume they think about the recommender system evolving in such a way that it remains close to its previous state. Additionally, the average song characteristics in the industry do not rapidly shift from day to day, but do have long-term trends, so they fit a first-order Markov process well.

I define the rightsholder expected profit function:

$$E[\pi_j(X_j, \mathcal{X}, \phi)] = 0.6\left(\sum_{t=0}^{T} \delta^t R_t \left(\frac{P(i \text{ listens to } j \text{ with characteristics } X)}{\sum_K P(i \text{ listens to } k \text{ with characteristics } \mathcal{X})}\right)\right) - F_j(X_j) \tag{11}$$

Each period t, defined as a day, the rightsholder owning song j receive a share of Spotify's gross revenue R_t. I assume that Spotify's revenue is exogenous, and that rightsholders know how much revenue Spotify earns each period. The first term on the right-hand side is the firm's variable profit, or the expected revenue from the song, which I subsequently denote VP_j Note that the rightsholder's variable profits do not include marginal costs. As digital goods, the marginal cost of releasing to a label of consumers streaming a song is effectively zero. I further discuss the functional form of the probability of listening to a song in section 5, including how the Markov processes enter into this probability.

4.4.2 Fixed Costs

Intuitively, rightsholders will not perfectly know the costs of writing, producing, and releasing a song beforehand, but they will possess experience-based knowledge about the typical costs and how they might vary with song characteristics. For example, a highly instrumental track might incur different production costs compared to a song requiring multiple vocalists and complex lyric writing. Therefore, I model these fixed costs as independently and identically distributed (i.i.d.) across songs following a lognormal distribution. This distribution is suitable as costs must be positive, and it allows for a flexible shape, potentially capturing a long tail of high-cost songs. The parameters of the distribution (location μ_j and scale σ_j) are allowed to vary as a function of the song's characteristics X_j.[24]

Formally, I define F_j as follows:

$$F_j \sim lognormal(\mu_j(X_j), \sigma_j(X_j)) \tag{12}$$

$$\mu_j(X_j) = \mu_g + X_j'\beta_\mu \tag{13}$$

$$\sigma_j(X_j) = \exp(\gamma_g + X_j'\beta_\sigma) \tag{14}$$

Here, μ_g and γ_g are global parameters representing the aggregate location and scale parameters. β_μ and β_σ are vectors of song characteristic coefficients capturing the effects of covariates on the location and scale, respectively. The exponential parameterization of the scale ensures that $\sigma_j > 0$ for all songs.

Specifically, their expected profit from releasing the song must be nonnegative. If the expected revenue exceeds fixed cost, the rightsholder releases the song; otherwise, it does not. This entry condition provides the upper bound to the fixed cost of releasing a song.

4.5 Equilibrium

4.5.1 Theoretical Framework: Oblivious Equilibrium

My solution concept is an oblivious equilibrium (Weintraub, Benkard, and Van Roy 2008), wherein market participants make optimal entry decisions based on simplified representations of the industry state. This equilibrium concept was developed to analyze dynamic oligopoly models with a large number of firms, where computing a standard Markov Perfect Nash Equilibrium (MPNE) would be computationally intractable. In an oblivious equilibrium, firms make decisions based only on their own state and the long-run average industry state, ignoring the specific states of their rivals. This dramatically reduces the effective state space from the full industry state to just the firm's own state and a fixed, long-run average distribution of rivals. The motivation behind this simplification is that in large markets, individual firm changes have negligible impact on the aggregate, and idiosyncratic shocks average out over time.

This equilibrium concept is especially appropriate for the music streaming industry, where thousands of new songs are released daily across diverse genres and styles. In such a vast market, no single song or rightsholder can significantly impact the aggregate industry state, and individual rightsholders realistically cannot track every competing song. Instead, they make strategic decisions based on their own song's characteristics and a general

24. I also considered a Weibull distribution, but the lognormal distribution provides a better fit to the data.

understanding of the competitive landscape, precisely the scenario oblivious equilibrium was designed to model.

In introducing such an equilibrium, Weintraub, Benkard, and Van Roy (2008) use an entry and investment game, where firms choose whether to enter the market, and whether to invest in improving their product quality x. In this game, firms face a fixed cost of entry F_j, and they receive a stream of profits $\pi(x, s)$, where x is the firm's quality and s is the industry state (quality of rival firms). They face a fixed cost to enter κ, an investment function $w(\iota, \zeta)$, where ι is the amount to invest and ζ is an investment shock, and a scrap value ϕ. The entry and investment game is a dynamic game, where firms must consider the future evolution of the industry state when making their decisions.

The assumptions necessary for such an oblivious equilibrium in such a game are the following:

1. $\forall x, s \in \mathcal{N},\ \pi(x, s)$ is increasing in x, the firm's quality, and decreasing in s, the strength of competition. It is also positive and bounded.

2. The random variables ϕ and ζ that affect decisions are independent and identically distributed (i.i.d.) conditional on state.

3. The investment function w is positive but bounded, and that investment transition functions are continuous.

4. The number of entering firms is a Poisson random variable whose entry rate λ is conditional on state. Additionally, $\kappa > \beta \cdot \bar{\phi}$, where β is the discount factor and $\bar{\phi}$ is the expected scrap value. This ensures that firms will not enter the market just to sell their scrap value.

In the context of the entry and investment game Weintraub, Benkard, and Van Roy 2008 discuss, and oblivious equilibrium consists of an investment strategy, defined as μ, and an entry rate, λ, such that:

1. Firms choose a strategy that maximizes their oblivious value function $\tilde{V}$

2. Either the expected value of entry is zero, or the entry rate $\lambda = 0$.

This approach has strong theoretical foundations. Weintraub, Benkard, and Van Roy (2008) show that as the market size grows large, the Oblivious Equilibrium strategies and outcomes approximate those of the true Markov Perfect Equilibrium. For an oblivious equilibrium to approximate a Markov Perfect Equilibrium, the equilibrium must satisfy a light-tail condition. This condition requires that states where a small change in the fraction of firms has a large effect on the value function must have a low probability in the long-run distribution of states.

This light-tail condition means that very large firms (or large levels of market concentration) are unlikely to occur in the long-run distribution of states. If this condition holds, then the oblivious equilibrium will approximate the Markov Perfect Equilibrium, as the optimal strategy under an oblivious value function will equal to the optimal strategy where firms track the true industry state, in expectation.

4.5.2 Model Equilibrium: Oblivious Entry in Music Streaming

Having described the theoretical framework of oblivious equilibrium, I now apply it to my model of the music industry. In my model, each rightsholder acts as an oblivious agent, choosing whether to release its song based on the song's characteristics, the long-run average characteristics of all songs (rather than tracking each rival's specific characteristics), and the probability the recommender system will recommend their song.

Each rightsholder has the following oblivious value function:

$$V_j(a_j, X_j, \mathcal{X}, \phi) = \max_{a \in \{0,1\}} \sum_{t=0}^{T} \delta^t E[VP_j(a_j, X_j, \mathcal{X}, \phi)] - F_j(a_j, X_j) \tag{15}$$

In the value function, each rightsholder has a one-time binary decision to release a song, a_j, and the value function is the variable profit VP_j from releasing the song, given the song's characteristics X_j, the average characteristics of all songs $\mathcal{X}$, and the probability the recommender system will recommend their song ϕ, minus the fixed cost of releasing the song $F_j(X_j)$ The profit function in the value function is given by equation 11, and the fixed cost of releasing the song is given by $F_j(a_j, X_j)$.

Next, I confirm that this model satisfies the conditions for oblivious equilibrium. First, the profit function $\pi(a_j, X_j, \mathcal{X}, \phi)$ is increasing in own-product characteristics X_j and decreasing in rival characteristics $\mathcal{X}$ through the choice probability (see equation 8). Second, the draw of fixed costs $F_j(x_j)$ is i.i.d. across songs, conditional on the song characteristics. Third, because this is a onetime release decision, no investment function is necessary. Fourth, instead of a Poisson process for entry, I limit my analysis to the number of songs released that enter Spotify's Top 200 each day, setting an upper bound on the number of songs that can enter the market. It is also not possible to exit the market, because music on a digital platform is a durable good.

I also confirm that the model satisfies the light-tail condition for oblivious equilibrium to approximate a Markov Perfect Equilibrium. With thousands of songs being released daily, no single song dominates the market to the extent that other rightsholders need to explicitly track its state. This approach is particularly appropriate given that rightsholders in the music industry rarely have perfect information about the exact characteristics of every competing song or the precise mechanisms of the recommender system.

Formally, in my model's oblivious equilibrium, rightsholders have an oblivious entry strategy a_j such that:

1. Rightsholder strategies maximize their oblivious value function: $V_j(a_j, X_j, \mathcal{X}, \phi)$, given consumer demand, the recommender system.

2. The Markov processes governing firm perception of the recommender system ϕ and rival songs $\mathcal{X}$ are stationary and consistent with the firm's optimal strategy.

3. Rightsholders only enter the market if their expected profit exceeds the fixed cost of releasing the song:

$$0.6\left(\sum_{t=0}^{T}\delta^t R_t\left(\frac{P(i \text{ listens to } j \text{ with characteristics } X)}{\sum_K P(i \text{ listens to } k \text{ with characteristics } \mathcal{X})}\right)\right) \geq F_j(X_j) \tag{16}$$

This closes the model, which I take to the data described in chapter 3.

5 Estimation

My estimation strategy has several stages:

1. Demand and Recommender System estimation
2. Markov Process estimation
3. Expected revenue calculation
4. Fixed cost estimation

5.1 Consumer Demand Estimation

In the first stage, I estimate consumer preferences and recommender system preferences using the MSSD data. I separately estimate for consumers who receive songs from the recommender system, and those who do not. Specifically, I estimate $\theta_1 = (\beta, \gamma, \eta)$ from equations 4 and 7 in this stage. These characteristics include song characteristics, consumer characteristics, and session fixed effects. The song and consumer characteristics are the ones detailed in section 1 and 2, and the session fixed effects are the position of the song in the streaming session. For consumer preferences, I use the Fox et al. (2011) (FKRB) estimator, which creates a grid of consumer types based on their heterogenous preferences β_i and assigns a type weight w_i to each type. As a simple example, consider a random coefficient β_i on a song characteristic, say tempo, with as assumed support from $[-1, 1]$. The FKRB estimator creates a grid of types (coefficents) and estimates the weight of each type using OLS. In this example, the estimating equation is:

$$P(i \text{ listens to } j) = \sum_i w_i \frac{\exp(\beta_{r,i} X_j)}{1 + (\exp(\beta_{r,i} X_j) + \exp(\beta_{r,i} E_{is}[X_j | X_{j,s-1}]))}$$

where w_i is the weight of type i, and $\beta_{r,i}$ is the coefficient of type i on the song characteristic, tempo in this example.

To address the selection of grid points, I use the location-scale variant of this model, where I model each random coefficient β_i as a function of a mean β_μ and a scale parameter β_σ, such that $\beta_i = \beta_\mu + \beta_\sigma \cdot D$, where $D \in [0, 1]$ is a grid of points. This approach allows me to estimate the mean and variance of the random coefficients, rather than having to choose specific grid points. The estimating equation is now:

$$P(i \text{ listens to } j) = \sum_i w_i \frac{\exp((\beta_\mu + \beta_\sigma \cdot D_i) X_j)}{1 + (\exp((\beta_\mu + \beta_\sigma \cdot D_i) X_j) + \exp((\beta_\mu + \beta_\sigma \cdot D_i) E_{is}[X_j | X_{j,s-1}]))}$$

Now, I use numerical optimization to estimate β_μ and β_σ and w_i simultaneously.

For non-random coefficients, I fix their scale parameter to zero, so that they do not vary across types. Specifically, I allow six coefficients to be random: the alternative-specific constants for listen and skip, duration and its quadratic term, and tempo and its quadratic term. All other song characteristic coefficients are treated as non-random (homogeneous across consumers). I estimate this model in two steps: first, for a given set of β_μ and β_σ, I estimate the type weights w_i using OLS. Then, I use numerical optimization to find the β_μ and β_σ that maximize the likelihood of the observed choices, given the estimated type weights. I repeat this process until convergence, where the change in the likelihood function is below a certain threshold.

I identify my parameters through variation in the choices each consumer faces at each position in the streaming session. Because consumers received different songs at different positions in their streaming session, and their skip value adaptively evolves based on their history, each observation has significant variation in the characteristics space. This variation allows me to identify the parameters of the model.

I estimate the recommender system parameters using a maximum likelihood estimator over the probability a consumer completes a song. It uses the same song and consumer characteristics as the demand model, but does not include session fixed effects and has a different outcome variable (completion rather than 30-second listen).

5.2 Markov Process Estimation

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 in Spotify's top 200 to a consumer. I then compute the average of these probabilities across all songs in the Top 200 each day. This object is the average probability the recommender system surfaces a song to a consumer, which I denote as ϕ_t. I then estimate a SARIMAX model for ψ_0 and ψ_1. I incorporate one lag and a drift term to capture the long-term trend in the recommender system. For song characteristics, I compute the average characteristics of all songs on Spotify's Top 200 each day, and I estimate ν_0 and ν_1 as a Vector Autoregression (VAR) model. I use a VAR(1) model, because the average characteristics of songs on Spotify's Top 200 do not change rapidly from day to day, but do have long-term trends. I also include a drift term to capture the long-term trend in the average characteristics of songs on Spotify's Top 200.

5.3 Expected Revenue Calculation

In the third stage, I compute the expected revenue for each song released between January 1, 2018, and September 30, 2018. This timeframe aligns with the MSSD data period used for demand and recommender system estimation, ensuring consistency across model components.

For each song, I construct a discounted stream of future revenues based on equation 11, applying the estimated demand parameters (θ_1) and Markov process parameters (θ_2) to predict how market characteristics and algorithmic exposure evolve over time. I use a daily discount factor of $\delta = 0.9978$, corresponding to an monthly discount rate of approximately 93%, following Lee 2013 The calculation accounts for both immediate post-release performance and the song's long-term revenue stream over a three-year horizon (1,095 days), which typically exhibits decay patterns as shown in Figure 6.

In computing the future revenues, I define the firm's streamshare as the sum of two channels: an algorithmic (RS) channel, where the song is surfaced by the recommender system, and a direct discovery channel, where consumers find the song through non-algorithmic means. For each channel, I use the mixed logit demand model to compute the ratio of the song's predicted listens to total predicted listens across all competing songs:

$$s_{jt} = \underbrace{\frac{\phi_{jt} \exp(\beta^r X_j)}{\phi_{jt} \exp(\beta^r X_j) + N \cdot \bar{\phi}_t \exp(\beta^r \bar{X}_t)}}_{\text{RS channel}} + \underbrace{\frac{(1-\phi_{jt}) \exp(\beta^r X_j)}{(1-\phi_{jt}) \exp(\beta^r X_j) + N \cdot (1-\bar{\phi}_t) \exp(\beta^r \bar{X}_t)}}_{\text{Direct discovery channel}}$$

Here, ϕ_{jt} is the predicted probability the recommender system surfaces song j at time t (from the RS logistic model), $\bar{\phi}_t$ is the average rival RS probability (from the SARIMAX forecast), β^r is the random coefficient vector for consumer type r (from the FKRB grid), X_j are song j's characteristics, $\bar{X}_t$ are the mean characteristics of rival songs at time t (from the VAR forecast), and $N = 3{,}759$ is the effective number of rival songs.

The streamshare is then integrated over the FKRB mixture distribution by weighting across all consumer types:

$$\bar{s}_{jt} = \sum_{r=1}^{R} \theta_r \cdot s_{jt}^r$$

where θ_r are the estimated mixture weights from the FKRB estimator.

To account for the fact that consumers cannot discover songs through non-algorithmic channels immediately upon release, I set the direct discovery channel to zero for the first 14 days after release. This reflects the time required for organic discovery through radio, word-of-mouth, social media, and editorial playlist placement.[25]

The effective number of rival songs, $N = 3{,}759$, is calibrated to match the observed mean streamshare in the data. I use the observed daily mean characteristics of songs on Spotify's Top 200 as the initial rival state $\bar{X}_0$, looked up by each song's release date. These rival characteristics and RS probabilities evolve stochastically over the three-year horizon through the estimated VAR(1) and SARIMAX(1,0,0) processes, respectively. I average over multiple Monte Carlo draws of these stochastic processes to compute the expected streamshare at each horizon.

The daily revenue each song generates is its streamshare times the rightsholder's share (60%) of Spotify's daily premium subscription revenue. I use Spotify's quarterly earnings reports to construct a daily revenue series from 2017 through 2024. Expected revenue is the discounted sum of daily revenues over the three-year horizon:

25. Spotify Support

$$E[VP_j] = \sum_{t=0}^{1095} \delta^t \cdot R_t \cdot \bar{s}_{jt}$$

where R_t is the rightsholder's daily revenue. When computing expected revenues, I standardize all audio features using the MSSD sample means and standard deviations, and winsorize them at ± 3 standard deviations to prevent extreme values from generating implausible revenue predictions through the quadratic terms. These expected revenue estimates form the foundation for the subsequent fixed cost estimation.

5.4 Fixed Cost Estimation

The fourth stage estimates the distribution of fixed costs associated with releasing songs onto Spotify's platform. This estimation is critical for understanding entry decisions by rightsholders and represents a key contribution of this research.

I model fixed costs as following a lognormal distribution, which has several desirable properties for this application. First, it constrains costs to be positive, aligning with economic reality. Second, it allows for right-skewed cost distributions, which matches industry accounts of occasional high-budget productions. Third, its parameters have clear economic interpretations as location and scale.

The empirical challenge is that fixed costs are not directly observed. I address this using a revealed preference approach: if a song entered the market, its expected revenue must have at least covered its fixed cost. To operationalize this insight, I transform the expected revenue estimates from the previous stage into fixed cost observations by scaling them with industry-specific gross profit margins.

Specifically, I utilize Earnings Before Interest, Taxes, Depreciation, and Amortization (EBITDA) margins for the major labels to scale expected revenue for their songs.[26] For songs from independent labels, I apply a standard 20% EBITDA margin based on industry benchmarks, because I lack direct data on these margins. This approach accounts for the different profitability structures across label types while ensuring that the revealed fixed costs are economically meaningful. Using EBITDA margins is particularly useful in this context, because the goods are digital and have negligible marginal costs. As such, the EBITDA margin is a good proxy for the fixed costs of producing and releasing a song. It is, however, a measure of accounting cost, so this approach results in a lower bound estimate of fixed costs, as it does not directly include the opportunity costs of the rightsholder's

26. Music Business Worldwide, Sony Music Earnings Report, Music Business Worldwide, UMG Earnings Report, Music Business Worldwide, Warner Music Earnings Report

time and effort. This may bias my fixed cost estimates downwards, but it is a reasonable assumption given the lack of data on these opportunity costs.

Table 3 reports the EBITDA margins for the major labels:

Label	EBITDA Margin
Sony	24.4%
Universal	16.5%
Warner	19.6%

Table 3: EBITDA Margins for Major Labels

To obtain the fixed cost estimates, I use the following transformation:

$$F_j(X_j) = E[VP_j(X_j, \mathcal{X}, \phi)] \cdot \left(1 - \text{EBITDA Margin}\right) \tag{17}$$

The log-likelihood function for a single observation is:

$$\ell_j(\theta) = -\frac{1}{2}\log(2\pi\sigma_j^2) - \frac{(\log(F_j) - \mu_j)^2}{2\sigma_j^2} \tag{18}$$

The full log-likelihood aggregates across all observations:

$$L(\theta) = \sum_{i=1}^{n} \ell_j(\theta) \tag{19}$$

I maximize this function using the BFGS algorithm with numerical gradients. One limitation of this approach is that it only identifies the distribution of fixed costs for songs that successfully entered the market (i.e., appeared in Spotify's Top 200). This potential selection bias is mitigated by the fact that my sample includes a wide range of commercial success levels, from major hits to songs that briefly appeared in the charts before exiting. Additionally, this model allows me to extrapolate to song types underrepresented in the observed data.

This estimation results in a distribution of fixed costs for each song. When reporting results and in computing the conunterfactuals, I use the median fixed cost for each song, which will vary by song characteristics. This median fixed cost is the value of $F_j(X_j)$ in equation 16.

6 Results

Tables 4 and 5 reports consumer demand estimates for both direct selection and recommender system selection:

	Dependent variable: Royalty Bearing Stream (RBS)					
	Direct Selection			Recommender System Selection		
	Location	Odds Ratio	Scale	Location	Odds Ratio	Scale
Acousticness	0.0058	1.001		-0.0127	0.999	
Acousticness^2	-0.0049	–		0.0121	–	
Danceability	0.1880	1.105		0.2085	1.034	
Danceability^2	-0.0877	–		-0.1752	–	
Duration	0.0192	0.984	0.0539	-0.0124	0.970	0.0487
Duration^2	-0.0358	–	0.0851	-0.0183	–	0.0919
Energy	0.2693	1.084		-0.0377	0.968	
Energy^2	-0.1888	–		0.0053	–	
Instrumentalness	-0.0146	0.997		-0.1405	1.001	
Instrumentalness^2	0.0114	–		0.1419	–	
Liveness	0.0394	1.000		0.0136	0.992	
Liveness^2	-0.0396	–		-0.0212	–	
Loudness	0.0326	1.035		-0.0002	0.992	
Loudness^2	0.0023	–		-0.0078	–	
Speechiness	0.0654	1.010		0.1113	1.020	
Speechiness^2	-0.0553	–		-0.0919	–	
Tempo	0.1754	1.084	0.1356	0.7964	1.394	0.1239
Tempo^2	-0.0943	–	0.0457	-0.4640	–	0.0479
Valence	0.1675	1.040		0.1054	1.034	
Valence^2	-0.1285	–		-0.0719	–	
Age	-0.0276	0.973		-0.0096	0.990	
Mode	0.0120	1.012		0.0099	1.010	
Time Signature	0.0044	1.004		0.0386	1.039	
Model Statistics						
Observations	4,314,888			733,848		
Log-Likelihood	-1,129,147.5			-183,542.5		
Estimator	FKRB Location-Scale			FKRB Location-Scale		
Grid Points	600			600		

Notes:
Location and Scale are FKRB location-scale parameters (β_i = Location + Scale $\cdot$ D).
Odds ratios computed at the mean consumer type. Scale is blank for homogeneous coefficients.

Table 4: Consumer Demand Estimates – Song Characteristics

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

I focus on the odds ratios of the coefficients, which are more interpretable than the raw coefficients. The odds ratios represent the multiplicative change in the odds of a song being completed for a one-unit increase in the characteristic. For example, the odds ratio on danceability in direct selection is 1.105, which means that a one standard deviation increase in danceability increases the odds of a song being an RBS by 10.5%.

The results reveal striking differences between how users interact with music through direct selection versus recommender systems. In direct selection, song characteristics have modest but positive effects for danceability (odds ratio 1.105), energy (1.084), and tempo (1.084). Age has a small negative effect on direct demand, suggesting users slightly prefer newer songs when choosing directly. Interestingly, acousticness and valence (emotional

	Dependent variable: Royalty Bearing Stream (RBS)			
	Direct Selection		Recommender System Selection	
	LISTEN	SKIP	LISTEN	SKIP
Time of Day				
Morning	0.0557	-0.0206	0.0295	-0.0778
Day of Week				
Weekday	-0.0667	-0.0692	0.0953	0.0860
User Characteristics				
Premium	-0.2074	-0.2444	0.1082	0.0184
Session				
Log Position	0.0454	0.2910	0.1145	0.2901
Alternative-Specific Constants				
ASC (Listen) (Location)	2.4163		2.2924	
ASC (Listen) (Scale)	0.2274		0.2327	
ASC (Skip) (Location)	1.5497		1.8052	
ASC (Skip) (Scale)	0.2951		0.3039	
Model Statistics				
Observations	4,314,888		733,848	
Log-Likelihood	-1,129,147.5		-183,542.5	

Table 5: Consumer Demand Estimates – Contextual Characteristics

FKRB location-scale estimates. LISTEN and SKIP columns show alternative-specific coefficients.

positivity) show positive linear terms but negative quadratic terms in direct selection, indicating users prefer moderate levels of these characteristics when actively choosing music.

The recommender system demand parameters show a different pattern. Most strikingly, tempo has a much stronger positive effect in recommended songs (odds ratio 1.394) compared to its modest effect in direct demand (1.084), suggesting the recommender system effectively surfaces high-tempo songs that users enjoy. Danceability remains modestly positive (odds ratio 1.034), while energy shows a slightly negative effect (0.968) under recommender system selection, in contrast to its positive effect in direct selection.

The contextual characteristics also reveal interesting patterns. Premium users are less likely to engage with directly chosen songs but more likely to engage with recommended songs, suggesting they may be more trusting of or receptive to recommendations. They might also be more passive listeners, preferring to let the recommender system guide their listening. Time of day effects are stronger for both types of demand during morning and afternoon hours, with slightly larger coefficients for recommended songs. Day of week effects show modest variation, with weekday effects differing in direction between direct selection and recommender system contexts.

Tables 6 and 7 reports the results for the recommender system:[27]

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

	Dependent variable: Song Completion	
	Estimate (Std. Error)	Odds Ratio
Acousticness	0.0516*** (0.0056)	1.042
Acousticness^2	-0.0103*** (0.0017)	–
Danceability	-0.0275*** (0.0052)	0.975
Danceability^2	0.0025* (0.0013)	–
Duration	-0.1887*** (0.0059)	0.826
Duration^2	-0.0025*** (0.0008)	–
Energy	-0.0712*** (0.0069)	0.930
Energy^2	-0.0018 (0.0015)	–
Instrumentalness	-0.0539*** (0.0074)	0.952
Instrumentalness^2	0.0049*** (0.0007)	–
Liveness	-0.0279*** (0.0043)	0.973
Liveness^2	0.0009 (0.0010)	–
Loudness	0.0658*** (0.0056)	1.067
Loudness^2	-0.0009 (0.0006)	–
Speechiness	-0.0352*** (0.0042)	0.970
Speechiness^2	0.0046*** (0.0010)	–
Tempo	-0.0351*** (0.0053)	0.975
Tempo^2	0.0098*** (0.0016)	–
Valence	0.0151*** (0.0051)	0.998
Valence^2	-0.0169*** (0.0015)	–
Age	-0.0942*** (0.0020)	0.910
Mode	-0.0010 (0.0034)	0.999
Time Signature	0.0337*** (0.0123)	1.034
ML Product Interaction	-0.0003** (0.0002)	1.000
CF Dot Product	0.0002*** (0.0000)	1.000
Context Switch	0.3976*** (0.0079)	1.488
Model Statistics		
Observations	1,682,912	
Pseudo R^2	0.007734	

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

Table 6: Recommender System Estimates – Song Characteristics

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

	Dependent variable: Song Completion	
	Estimate (Std. Error)	Odds Ratio
Time of Day		
Morning	0.1371***	1.147
	(0.0046)	
Afternoon	0.0771***	1.080
	(0.0040)	
Night	0.1409***	1.151
	(0.0066)	
Day of Week		
Tuesday	0.0046	1.005
	(0.0060)	
Wednesday	0.0202***	1.020
	(0.0061)	
Thursday	-0.0214***	0.979
	(0.0062)	
Friday	-0.0098*	0.990
	(0.0060)	
Saturday	0.0008	1.001
	(0.0062)	
Sunday	0.0079	1.008
	(0.0062)	
User Characteristics		
Premium	-0.0743***	0.928
	(0.0042)	
Intercept	-0.7812***	
	(0.0139)	
Model Statistics		
Observations	1,682,912	
Pseudo R^2	0.007734	

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

Table 7: Recommender System Estimates – Contextual Characteristics

Premium consumers are less likely to complete songs.

The recommender system's song selection behavior differs notably from how users engage with its recommendations. While both models show that age negatively influences recommendations and consumption, the effect is much stronger in the recommender system model (coefficient -0.0942 vs -0.0276 in direct demand), suggesting the system strongly favors newer songs. For song duration, the system shows a strongly negative preference (odds ratio 0.826), indicating it systematically surfaces shorter songs—a key mechanism by which the recommender system may have incentivized shorter music production. Similarly, the system shows negative effects for energy (odds ratio 0.930), danceability (0.975), and speechiness (0.970), suggesting it may be under-recommending songs with these characteristics relative to consumer preferences.

The mismatch between what consumers enjoy and what the system surfaces is particularly notable for energy and tempo. While the demand model showed users positively engage with energetic songs (odds ratio 1.084 in direct selection), the recommender system penalizes energy (odds ratio 0.930). Conversely, loudness shows a strong positive effect in the recommender system (odds ratio 1.067) that is not as prominent in direct demand. The quadratic terms for many characteristics also differ between the models, indicating that the system's understanding of optimal levels for these features does not perfectly align with what drives user engagement.

The contextual effects align with the consumer demand model in time-of-day patterns but differ on day-of-week patterns. The time-of-day patterns are similar, with both demand and consumption showing higher activity during morning and night hours. However, the day-of-week patterns differ notably: while the demand model showed higher engagement during midweek, the recommender system model shows consistently negative coefficients for all days relative to Monday. For premium users, the recommender system shows a negative effect (odds ratio 0.928), consistent with the pattern that premium subscribers are more selective listeners. Intuitively, premium subscribers, facing no ad interruptions, may be more likely to skip songs and search, whereas ad-supported users would prefer to avoid ads, and take a more passive approach to listening. The low pseudo-R^2 for the recommender system model (0.008) implies that these observable characteristics explain relatively little of the system's behavior, consistent with the system relying heavily on collaborative filtering and latent features not captured by the audio characteristics I observe.

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

This VAR reports 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

Statistic	Acousticness	Age	Danceability	Duration	Energy	Instrumentalness	Liveness	Loudness	Mode	Speechiness	Tempo	Time Signature	Valence
Constant	0.047	0.012	0.063*	0.065**	0.074	0.007	0.039	-0.045	0.031	0.001	0.063***	0.054	0.091
	(0.107)	(0.269)	(0.030)	(0.021)	(0.039)	(0.100)	(0.040)	(0.057)	(0.041)	(0.051)	(0.013)	(0.037)	(0.059)
Drift	0.000**	0.000	0.000	-0.000***	-0.000*	0.000	0.000	-0.000*	-0.000	0.000	0.000***	-0.000*	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Own Lag	0.922***	0.901***	0.904***	0.901***	0.905***	0.923***	0.905***	0.905***	0.908***	0.915***	0.900***	0.919***	0.902***
	(0.017)	(0.022)	(0.014)	(0.010)	(0.021)	(0.013)	(0.010)	(0.023)	(0.008)	(0.011)	(0.009)	(0.006)	(0.010)
Observations	1,825	1,825	1,825	1,825	1,825	1,825	1,825	1,825	1,825	1,825	1,825	1,825	1,825

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

Table 8: VAR Estimation – Song Characteristics

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

	Dependent variable: Predicted Probability ($\hat{\phi}_t$)
$\hat{\phi}_{t-1}$	0.898***
	(0.010)
Drift	0.000
	(0.000)
Constant	0.029***
	(0.000)
Observations	1,826
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 9: Markov Process Estimation for Recommender System

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

zero, further supporting my argument 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:

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

Figure 9 plots the distribution of expected revenue for songs released in 2018 that entered Spotify's top 200 at least once:

These songs have an expected revenue ranging from $537,000 to $233 million, with a mean of $3.5 million. The distribution is right-skewed, with a small number of hit songs generating very large revenues.

My fixed cost estimation produces estimates of the location (μ) and (σ) parameters of the lognormal fixed cost distribution as a function of song characteristics.

Table 10 reports the results of the fixed cost estimation:

The location model results reveal several key determinants of fixed costs in music production. Duration shows a significant negative linear effect (-0.320), indicating that longer songs generally have lower fixed costs, potentially reflecting that shorter songs require more intensive production to achieve commercial viability. Among audio features, liveness has a significant negative effect (-0.050), while speechiness has a positive effect (0.073). There are also notable differences across record labels—Universal songs show

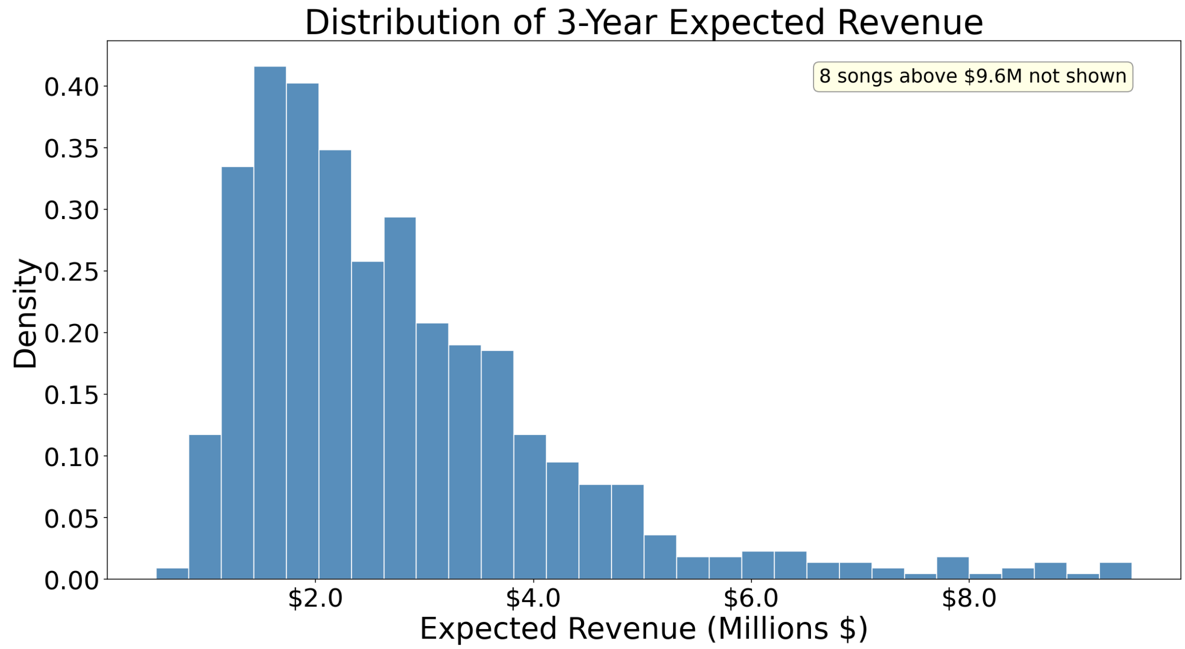

Figure 9: Expected Revenue of Songs Released in 2018 that Entered Spotify's Top 200

These songs have a mean expected revenue of $3.5 million.

significantly higher average fixed costs (coefficient 0.166), which could reflect different production strategies or accounting practices across labels.

The scale model results reveal interesting patterns in cost variability. Speechiness and acousticness are associated with higher cost variability (0.081 and 0.077), while duration shows a U-shaped relationship with cost variability, with the negative linear term (-0.432) and positive quadratic term (0.287) indicating that medium-length songs have the most predictable costs.

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

Figure 10 plots the distribution of median fixed costs predicted by the model:

The fixed cost of releasing a song on Spotify ranges from $1.2 million to $12 million, with a median of $1.9 million.

My estimated median fixed cost for songs in the top 200 is in line with a report from Chace (2011), which estimated the total cost of releasing Rihanna's "Man Down" at $1.2 million in 2011 dollars ($1.4 million in 2018 dollars), including $78,000 in production costs, $1 million in marketing, and $125,000 for the music video.

	Location (μ) Estimate (Std. Error)	Scale (σ) Estimate (Std. Error)
Intercept	7.422***	-0.856***
	(0.052)	(0.088)
Audio Features		
Acousticness	-0.046**	0.077***
	(0.015)	(0.023)
Instrumentalness	0.018	0.061
	(0.031)	(0.037)
Liveness	-0.050**	0.011
	(0.016)	(0.026)
Speechiness	0.073***	0.081**
	(0.017)	(0.025)
Loudness	0.022	0.067
	(0.030)	(0.043)
Duration		
Duration	-0.320**	-0.432**
	(0.106)	(0.134)
Duration^2	0.498***	0.287***
	(0.069)	(0.073)
Record Label		
Sony	0.083	0.126
	(0.052)	(0.092)
Universal	0.166***	0.139
	(0.045)	(0.079)
Warner	0.099	0.034
	(0.052)	(0.093)
Model Statistics		
Observations	749	
$\bar{R}^2$	0.323	–

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

Table 10: Fixed Cost Parameter Estimates

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

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 af-

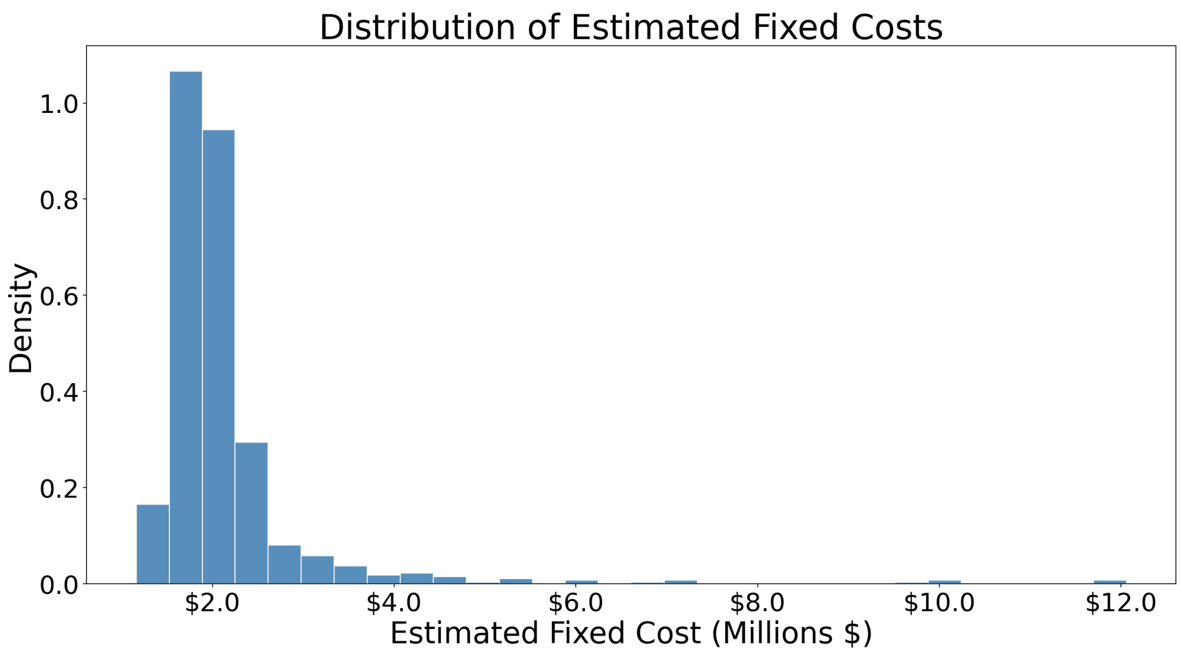

Figure 10: Distribution of Estimated Fixed Costs

These songs have a median fixed cost of $1.9 million.

fected the kind of music record labels are releasing. To isolate the impact of recommender systems specifically, I conduct two 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.

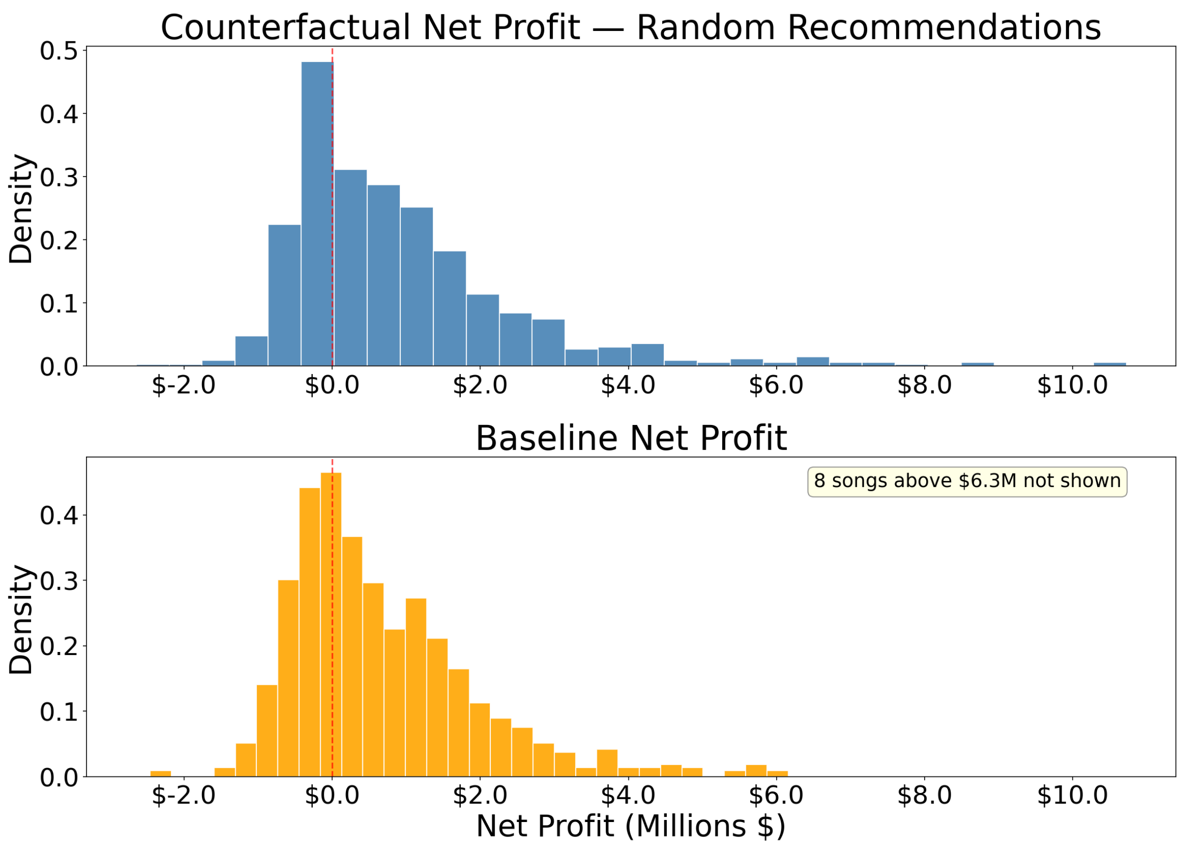

Figure 11: Counterfactual Expected Profit — Random Recommendations

Many songs become unprofitable when random recommendations are used.

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

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

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

First, I compare average expected profit for songs. Figure 11 reports the results of this comparison:

Each observation in this figure represents a song released in the first three quarters of 2018. Note that some of the estimated net profits are negative, because the median fixed cost estimated by the parameters of the lognormal distribution is higher than the expected revenue. Intuitively, it is likely that the realized fixed cost for those songs is lower than the median fixed cost. Many songs become unprofitable when random recommendations

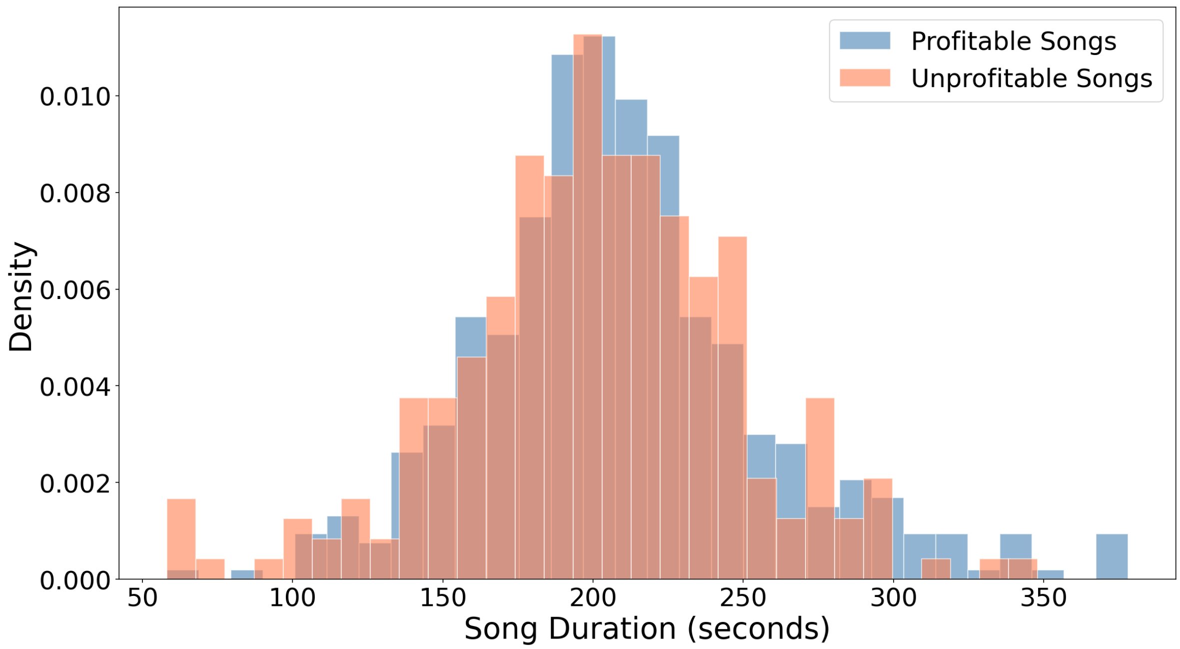

Figure 12: Counterfactual Duration — Random Recommendations

Unprofitable songs are shorter and more homogeneous than profitable songs.

are used. Indeed, of the 749 songs I observe that were released in the first three quarters of 2018, 248 (33%) are unprofitable. For those songs that are profitable, their gross profit margin is 35.6%.

I now turn to some song characteristic results and welfare implications of my counterfactual analysis. Figure 12 reports the average duration of songs between profitable and unprofitable songs.

The average duration of songs of profitable songs is 210 seconds, and the average duration of unprofitable songs is 202 seconds. This difference in means is significant at the 5% level. However, the Kolmogorov-Smirnov test does not reject the null that the two duration distributions are identical ($p = 0.21$), suggesting the distributional shift is modest. This suggests that while recommender systems allow somewhat shorter songs to enter the market, the effect on the overall duration distribution is limited.

Song tempo represents the most striking difference between profitable and unprofitable songs. Figure 13 reports the distribution of tempo for profitable and unprofitable songs.

Profitable songs average 140 BPM compared to 94 BPM for unprofitable songs, and the KS test strongly rejects distributional equality ($p \approx 0$).

Table 11 reports the average values of other song characteristics for profitable and unprofitable songs, the difference in means, and the difference in distributions (as evaluated

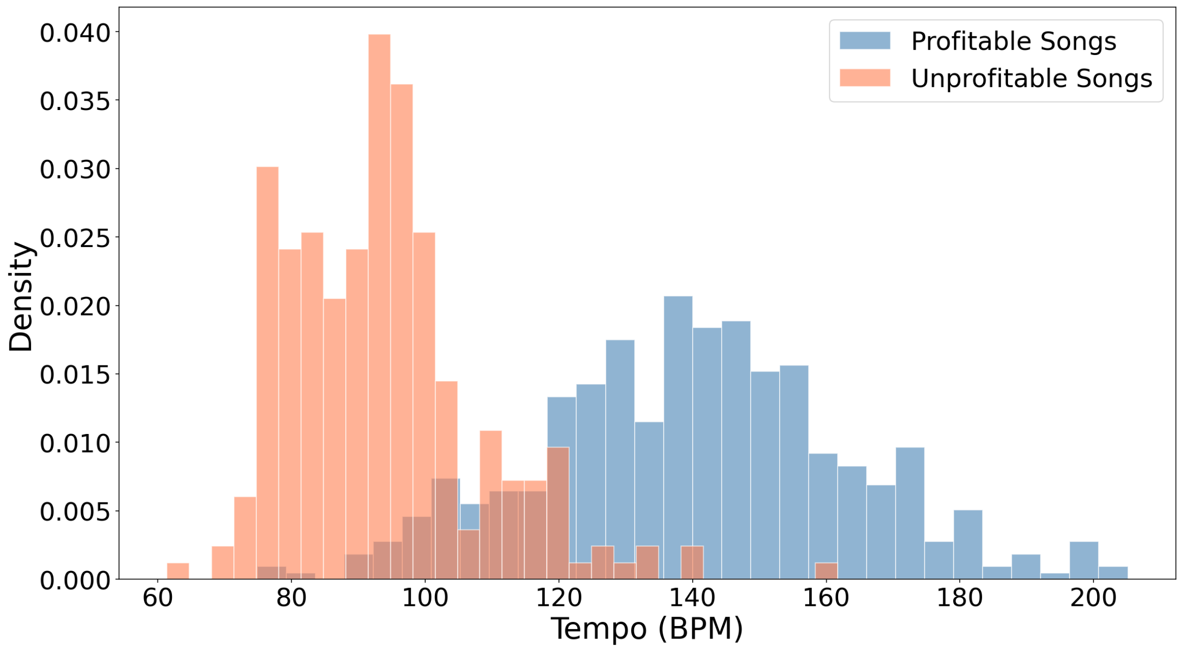

Figure 13: Counterfactual Tempo — Random Recommendations

Unprofitable songs have substantially lower tempo than profitable songs.

by a KS-Test):

The data reveals several striking differences between profitable and unprofitable songs' musical characteristics under random recommendations. Most notably, tempo differs dramatically between profitable and unprofitable songs, with profitable songs averaging 140 BPM compared to 94 BPM for unprofitable songs ($p \approx 0$). Profitable songs are also significantly longer, with an average duration of about 210 seconds compared to 202 seconds for unprofitable songs—an 8-second difference that is significant at the 5% level. Danceability also differs significantly ($p < 0.01$), with profitable songs showing higher values (0.713 vs. 0.675).

While some characteristics show differences in means, their distributional differences are less pronounced. Valence, energy, and mode show no significant distributional differences at the 5% level. Instrumentalness shows virtually identical distributions between profitable and unprofitable songs (KS p-value = 0.84), despite a small difference in means. Overall, this comparison suggests that the recommender system enables lower-tempo, shorter songs to enter the market that would otherwise be unprofitable.

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 expected log-sum of exponentiated utility, integrated over the estimated distribution of consumer preferences, following Anderson, Palma, and Thisse

Feature	Profitable	Unprofitable	Diff	KS-Test P-Value
Duration (s)	209.944	201.542	-8.402**	0.2105
	(2.150)	(3.028)		
Tempo	140.344	93.674	-46.670***	0.0000***
	(1.029)	(0.914)		
Energy	0.628	0.625	-0.003	0.1614
	(0.006)	(0.011)		
Danceability	0.713	0.675	-0.038***	0.0048***
	(0.006)	(0.010)		
Valence	0.441	0.434	-0.007	0.0592*
	(0.009)	(0.015)		
Acousticness	0.202	0.228	0.026	0.1090
	(0.010)	(0.016)		
Instrumentalness	0.008	0.009	0.001	0.8409
	(0.003)	(0.005)		
Liveness	0.172	0.186	0.014	0.4557
	(0.005)	(0.009)		
Speechiness	0.159	0.165	0.006	0.1119
	(0.006)	(0.010)		
Loudness	-6.315	-6.501	-0.185	0.6963
	(0.093)	(0.177)		
Mode	0.581	0.552	-0.028	0.9986
	(0.022)	(0.032)		

Table 11: Counterfactual Song Characteristics – Random Recommendations

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

(1992) and Train (2009). Because the demand model is a random-coefficients logit estimated by FKRB as a discrete mixture over a Halton grid of R support points $\{\beta^r\}_{r=1}^R$ with weights $\{\theta_r\}$, the exact expected consumer surplus under the mixed-logit demand is

$$\mathrm{E}[CS] = \sum_{r=1}^{R} \theta_r \log\left(\sum_{j=1}^{N} \exp\left(\beta^r X_j + \gamma_L Y_i + \eta_L \log(s)\right)\right), \tag{20}$$

where $\beta^r = a + b \odot z_r$ with a the FKRB location vector, b the scale vector, and z_r the r-th Halton point, and N represents the number of songs in the set, rather than the binomial skip-listen decision. This formulation is consistent with the mixed-logit demand model used to compute expected revenue, so welfare and entry incentives are evaluated under the same preference distribution. Integrating over consumer preference heterogeneity moderates the welfare cost of removing songs from the choice set, because consumers whose tastes are far from a removed song would be unlikely to choose it under the targeted system as well; this yields more conservative welfare estimates than the corresponding linear-logit log-sum evaluated at the mean coefficients alone. 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 1.9% higher when targeted recommender systems are used, compared to when random recommendations are used. Restated, random recommender systems result in a 1.9% decrease in consumer surplus. This suggests that recommender systems have increased consumer surplus by allowing for more songs to enter the market, and for consumers to find songs that they enjoy more easily.

7.2 Popular Recommendations

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

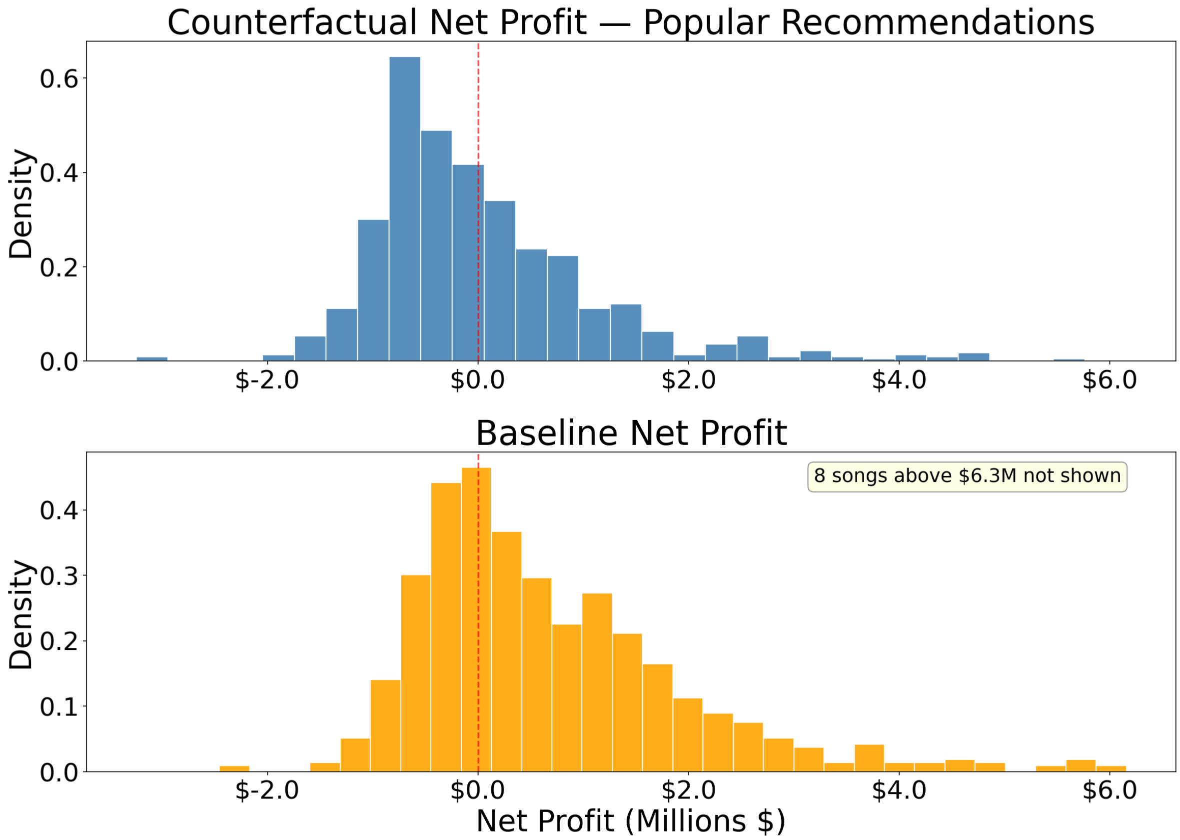

Figure 14: Counterfactual Expected Profit — Popular Recommendations

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

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 14 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 749 songs I observe that were released in the first three quarters of 2018, only 297 (40%) are profitable. For those songs that are profitable, their gross profit margin is 27.7%.

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

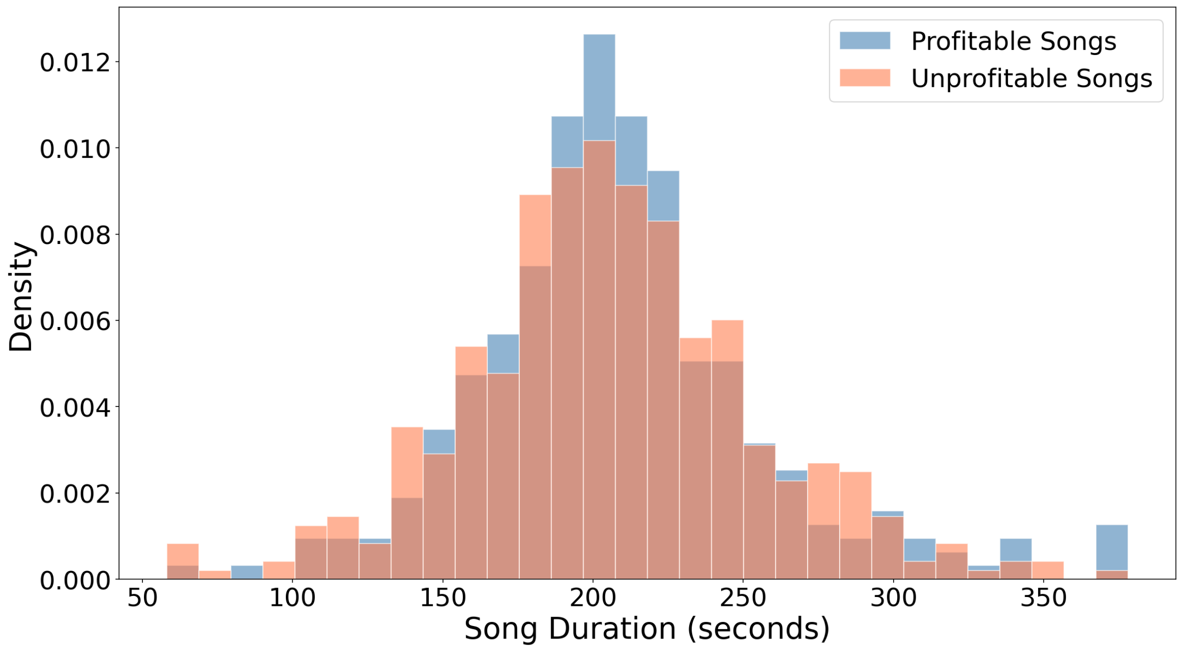

Figure 15: Counterfactual Duration — Popular Recommendations

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

The average duration of songs of profitable songs is 209 seconds, and the average duration of unprofitable songs is 206 seconds. This difference is not statistically significant.

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 a KS-Test):

The data shows more modest differences between profitable and unprofitable songs under popular recommendations. The most striking difference is in tempo, with profitable songs averaging 153 BPM compared to 107 BPM for unprofitable songs ($p \approx 0$), echoing the pattern observed under random recommendations. Duration shows a modest and statistically insignificant difference of about 3 seconds (209 vs. 206 seconds). Valence shows a marginally significant difference ($p = 0.018$), with profitable songs scoring 0.453 versus 0.428 for unprofitable songs. Speechiness also shows a significant distributional difference ($p = 0.002$), with profitable songs having slightly higher speechiness.

Notably, many musical characteristics show no significant differences in either means or distributions between profitable and unprofitable songs. Features such as energy, danceability, acousticness, instrumentalness, liveness, and loudness all have high KS test p-values, suggesting very similar distributions between the two groups. Overall, popular recommendations appear to primarily differentiate songs on tempo, consistent with the tempo-driven selection observed in the random counterfactual.

Feature	Profitable	Unprofitable	Diff	KS-Test P-Value
Duration (s)	209.010	205.948	-3.062	0.5242
	(2.759)	(2.280)		
Tempo	152.844	106.524	-46.319***	0.0000***
	(1.063)	(0.975)		
Energy	0.636	0.621	-0.015	0.1198
	(0.008)	(0.008)		
Danceability	0.699	0.701	0.003	0.7483
	(0.009)	(0.007)		
Valence	0.453	0.428	-0.025*	0.0184**
	(0.011)	(0.010)		
Acousticness	0.200	0.218	0.018	0.1077
	(0.013)	(0.011)		
Instrumentalness	0.008	0.008	-0.000	0.9754
	(0.003)	(0.003)		
Liveness	0.179	0.175	-0.004	0.1419
	(0.007)	(0.006)		
Speechiness	0.169	0.155	-0.013	0.0015***
	(0.007)	(0.007)		
Loudness	-6.256	-6.456	-0.200	0.2367
	(0.123)	(0.116)		
Mode	0.572	0.571	-0.002	1.0000
	(0.029)	(0.023)		

Table 12: Counterfactual Song Characteristics – Popular Recommendations

Popular recommendations generate a superstar effect, concentrating revenue among fewer songs.

Figure 16 plots the distribution of song valence for profitable and unprofitable songs:

Valence shows a marginally significant distributional difference under popular recommendations (KS $p = 0.018$), with profitable songs scoring higher (0.453 vs. 0.428).

Finally, I turn to the welfare implications of my counterfactual analysis. I find that consumer surplus is 5.3% higher when targeted recommender systems are used, compared to popular recommendations. Restated, popular recommender systems result in a 5.3% decrease in consumer surplus.

8 Conclusion

As recommender systems become increasingly embedded across digital platforms, understanding their impact on both consumer demand and producer decisions is crucial. This dissertation has developed a structural model of the music streaming industry to quantify

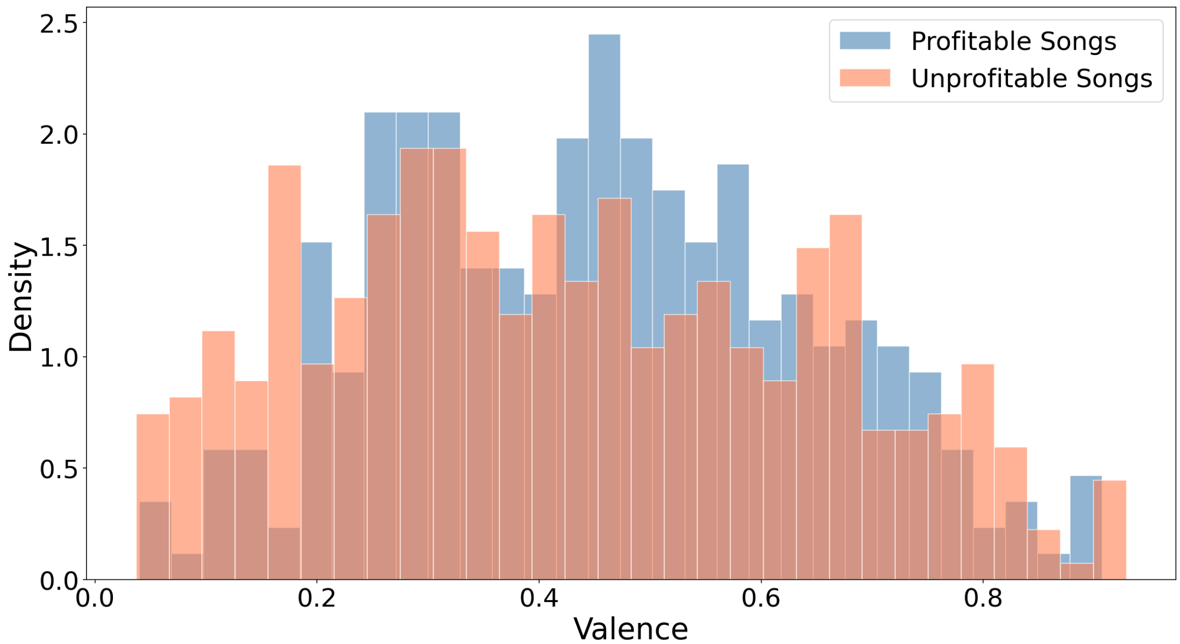

Figure 16: Counterfactual Valence — Popular Recommendations

Profitable songs are slightly more positive-sounding than unprofitable songs under popular recommendations.

how algorithmic recommendations influence music production and shape the evolving sound of popular music. The music industry has historically been at the forefront of technological disruption, making it an ideal setting to examine how recommender systems may affect other content-driven industries.

Using detailed data on streaming sessions and song characteristics from 2018, I find compelling evidence that recommender systems have fundamentally altered music production. The introduction of algorithmic recommendations appears correlated with incentives for the release of shorter, more sonically homogeneous songs optimized for platform objectives. These changes reflect a strategic response by producers to the recommender system's preferences rather than shifting consumer tastes, as demonstrated by my robustness analysis.

My counterfactual analyses reveal important welfare implications. While recommender systems may reduce musical diversity in some dimensions, they have simultaneously enabled more songs to profitably enter the market. I estimate that personalized recommendations increase consumer surplus by approximately 2% compared to random recommendations, suggesting that despite homogenization concerns, these systems create meaningful welfare gains by helping consumers discover music aligned with their preferences. A critical insight is the potential misalignment between recommender system objectives (like maximizing completed listens, as suggested by my recommender system model estimation) and broader consumer preferences or welfare metrics. This divergence stems partly from potentially conflicting incentives between platforms and rightsholders regarding how engagement is measured and rewarded. This dynamic also has complex implications for artists. While increased entry enabled by recommenders benefits some, the potential homogenization towards characteristics favored by the algorithm could disadvantage artists whose work deviates from these trends, potentially impacting discoverability and earnings, particularly for niche or independent creators. The popularity-based recommendation counterfactual further demonstrates that a system promoting only the most widely-consumed songs would generate a pronounced superstar effect, increasing profitability for a small subset of content while reducing overall consumer welfare significantly (by approximately 5% in the counterfactual) compared to personalized recommendations.

These findings have significant implications for platform regulation and competition policy. As policymakers scrutinize algorithmic curation, my results highlight both the efficiency gains from personalized recommendations and their potential to shape creative production and market structure. The framework developed here, including the fixed cost estimates of approximately $1.9 million per commercially viable song entering the Top 200, could inform antitrust analysis where algorithmic recommendations mediate between producers and consumers. Specific regulatory avenues warrant consideration, potentially including mandating greater transparency into recommendation algorithms' objectives and the data inputs used, exploring 'algorithmic disgorgement' or other remedies in antitrust cases involving biased recommendations, or establishing auditing frameworks to assess algorithmic impact on market diversity and creator inequality. Furthermore, the fixed cost estimates inform discussions concerning market entry barriers and could shape policies aimed at fostering a diverse and sustainable creator ecosystem.

For platforms, these results suggest opportunities to better align recommendation algorithms with both consumer preferences and business objectives. For producers, understanding how algorithms evaluate content characteristics provides strategic guidance for product design. The documented gap between what consumers enjoy and what recommender systems prioritize represents an opportunity for competitive differentiation.

Several promising avenues exist for future research. First, crucially, extending the empirical analysis with more recent data is needed to capture the rapid evolution of algorithms, user behavior (e.g., influenced by platforms like TikTok), and the music market since the 2018 dataset used here. Second, extending this model to other content platforms is a promising avenue. Contrasting findings from music streaming with platforms dominated by short-form, algorithm-driven user-generated content (like TikTok or YouTube Shorts) versus those with longer-form, professionally curated content (like Netflix) could yield critical insights into how algorithmic mediation differs across digital ecosystems. Third, incorporating random coefficients would enrich the consumer demand structure and better capture preference heterogeneity. Fourth, endogenizing platform decisions regarding pricing and recommendation algorithms would enable deeper analysis of platform market power and strategic manipulation of recommendations. Furthermore, examining the interplay between recommendation systems and the rise of generative AI in music creation presents a vital new research frontier: how will algorithms evaluate and surface AI-generated content, and will this accelerate or counteract the trends in characteristic optimization identified here? Finally, examining more directly how recommender systems affect market concentration and creator inequality remains an important question.

An emerging frontier for future research is the interaction between recommender systems and AI-generated music. As tools like Suno and Udio make music creation more accessible and potentially more algorithmically optimized from inception, the feedback loop between recommendation and production could become even tighter. This may accelerate the homogenization trends identified in this research, or paradoxically, enable greater diversity through lower production costs. Understanding this dynamic will be crucial as generative AI becomes more prominent in creative industries.

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

Appendix 1: Reduced Form Analysis

To movivate my structural model, I conduct reduced-form analysis of the relationship between song length and changes in music formats. Using songs that charted on Billboard's Hot 100 from 1940 to 2022, I estimate the following regression:

$$\begin{aligned} \text{Duration}_j &= \beta_0 + \beta_1 \mathbb{1}\{\text{Vinyl}\}_t + \beta_2 \mathbb{1}\{\text{Cassette}\}_t + \beta_3 \mathbb{1}\{\text{CD}\}_t \\ &= \beta_4 \mathbb{1}\{\text{Digital}\}_t + \beta_5 \mathbb{1}\{\text{Streaming}\}_t + \beta_6 \mathbb{1}\{\text{Recommenders}\}_t + \epsilon_j \end{aligned} \tag{21}$$

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 13 reports the results of this regression.

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

In addition to the above regression, I also conducted structural break tests to determine whether the introduction of streaming services and recommender systems caused a structural break in song length. Structural break tests are statistical procedures used to determine if the parameters of an economic or statistical model have changed significantly at some point (or points) in time within the sample period. A "structural break" or "structural change" implies that the underlying relationship between the variables in the model is not stable over the entire dataset. In the context of technology introduction, these tests are valuable because the adoption or implementation of a new technology can fundamentally alter economic relationships.

When the potential break date is known a priori, as is the case with the specific introduction dates of technologies, the Chow Test is the appropriate procedure. The Chow test formally compares the goodness-of-fit of a single regression model estimated over the entire sample period against the combined goodness-of-fit of separate regression models estimated on the sub-samples before and after the hypothesized break date. Specifically, it tests the null hypothesis that the coefficients in the regression model are the same across the different sub-periods against the alternative hypothesis that at least one coefficient differs.

To investigate whether the introduction of Spotify in 2011 and recommender systems

	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 13: Reduced Form Regression Results

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

in 2015 led to structural changes in the relationship described by the Billboard data, I conducted Chow tests at these two known break dates.

The regression model I used for the Chow test is the following:

$$\text{Duration}_{jt} = \beta_0 + \beta_1 \text{Year}_{jt} + \epsilon_{jt} \tag{22}$$

First, I tested for a structural break corresponding to the introduction of Spotify in 2011. The Chow F-statistic was calculated as 51.99, with an associated p-value of 0.00. Based on this result, I reject the null hypothesis of parameter stability at the 1% significance level for this first break point.

Second, I tested for a structural break corresponding to the introduction of recommender systems in 2015. The Chow F-statistic was calculated as 46.15, with an associated p-value of 0.00. Based on this result, I reject the null hypothesis of parameter stability at the 1% significance level for this second break point.

This analysis is consistent with anecdotal evidence of changes in songs since the introduction of streaming services, but it does not establish a causal relationship, or the mechanism by which these changes occur. For that I construct a structural model of the industry.

Appendix 2: 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:

$$\begin{aligned} log(\text{Streams}_{jt}) = \alpha + \beta_1 \text{Duration}_j + \beta_2 \text{Time Trend}_t \\ + \delta(\text{Duration} \times \text{Time Trend})_{jt} + \gamma X_j + \eta_t + \epsilon_{jt} \end{aligned} \tag{23}$$

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 14 reports the results of this regression:

	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 14: Difference-in-Differences Results

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

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

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

	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 15: Discrete Choice Model Results

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

Consumers have the following utility function:

$$\begin{aligned} U_{ijti} = \alpha + \beta_1 \text{Duration}_j + \\ + \delta(\text{Duration} \times \text{Time Trend}_{jt}) + \gamma X_j + \eta_t + \epsilon_{ijt} \end{aligned} \tag{24}$$

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 15 reports the results of this regression:

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

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

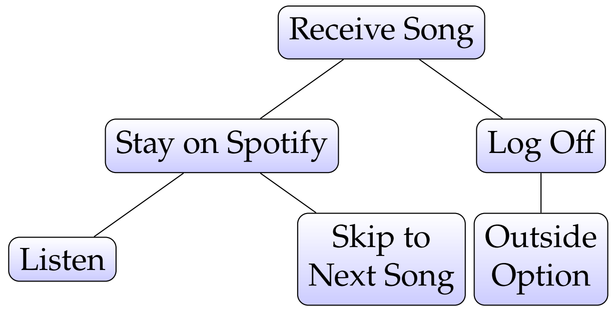

Figure 17: Nested Consumer Decision Tree

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

Appendix 3: Nested Logit Specification

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

I estimate a conditional logit and nested logit model to compare the two. Following Reimers and Waldfogel (2023), I estimate the nested logit model in a bottom-up fashion, estimating the inside options first, then the nest parameter.

Tables 16 and 17 report consumer demand estimates for both the conditional logit and nested logit structure:

The nesting parameters are $\lambda = 0.748$ for direct selection and $\lambda = 0.812$ for recommender system selection, indicating some substitution structure within the stay-on-platform nest. However, the song characteristic coefficients and odds ratios are very similar across the two specifications. For example, the danceability odds ratio is 1.105 in the conditional logit versus 1.106 in the nested logit for direct selection, and the tempo odds ratio is 1.084 versus 1.089. The largest differences appear in the recommender system subsample, where the tempo odds ratio increases from 1.394 to 1.424, but this difference is modest. Given the robustness of the coefficients to the nesting structure, I use the conditional logit specification in the main analysis for computational tractability.

	Direct Selection		Recommender System Selection	
Feature	Estimate	Odds Ratio	Estimate	Odds Ratio
Acousticness	0.0051	1.002	-0.0190	0.998
× Acousticness	-0.0033		0.0170	
Danceability	0.1989	1.106	0.2605	1.053
× Danceability	-0.0984		-0.2088	
Duration	0.0371	0.992	0.0207	0.982
× Duration	-0.0451		-0.0387	
Energy	0.2674	1.092	-0.0011	0.974
× Energy	-0.1791		-0.0248	
Instrumentalness	-0.0124	0.998	-0.1285	1.001
× Instrumentalness	0.0099		0.1300	
Liveness	0.0317	0.999	0.0065	0.991
× Liveness	-0.0326		-0.0153	
Loudness	0.0192	1.020	0.0193	1.022
× Loudness	0.0002		0.0022	
Speechiness	0.0596	1.009	0.1039	1.020
× Speechiness	-0.0504		-0.0836	
Tempo	0.1985	1.089	0.8543	1.424
× Tempo	-0.1134		-0.5009	
Valence	0.1332	1.030	0.1036	1.034
× Valence	-0.1032		-0.0703	
Age	-0.0241	0.976	-0.0105	0.990
Mode	0.0106	1.011	0.0092	1.009
Time Signature	0.0071	1.007	0.0272	1.028
Nesting Parameter (λ)	0.7480		0.8117	
Model Statistics				
Observations	4,314,888		733,848	
Grid Points	600		600	
Estimator	FKRB Mixed Nested		FKRB Mixed Nested	

Table 16: Logit and Nested Logit Demand Estimates – Song Characteristics

Nested logit specification. The nesting parameter λ captures substitution patterns within the stay-on-platform nest.

Supplemental Tables

Table 18 reports the descriptive statistics for the Spotify Charts data, covering the top 200 songs in the US between 2017 and 2021.

	Direct Selection		RS Selection	
Feature	LISTEN	SKIP	LISTEN	SKIP
Time of Day				
Morning	–	–	–	–
Day of Week				
Weekday	–	–	–	–
User Characteristics				
Premium	-0.1889	-0.2185	0.1268	0.0459
Session				
Log Position	0.0918	0.3019	0.1742	0.3385
Alternative-Specific Constants				
Listen	2.3314	–	2.2199	–
Skip	1.5817	–	1.7520	–
Observations	4,314,888		733,848	

Table 17: Logit and Nested Logit Demand Estimates – Contextual Characteristics

Contextual parameters for the nested logit specification.

	Mean	Median	Standard Deviation	Min	Max
Duration (s)	203.27	199.32	54.28	30.13	943.53
Release Year	2017	2019	7	1942	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	-38.85	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	0.93	1.00	0.26	0.00	1.00
Valence	0.46	0.46	0.22	0.03	0.98

Table 18: Spotify Charts Song Characteristics ($N = 9,245$ songs)

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

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 19: Descriptions of Song Characteristics

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 20: Examples of Song Characteristics

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

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

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

Table 21: Full Sample Consumer Demand Estimates - Song Characteristics

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

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

Table 22: Full Sample Consumer Demand Estimates - Contextual Characteristics

	Acousticness	Age	Danceability	Duration	Energy	Instrumentalness	Liveness	Loudness	Mode	Speechiness	Tempo	Time Signature	Valence
Constant	0.447**	-0.325	-0.320*	0.028	-0.320**	-0.131***	-0.104	-0.015	0.061	-0.044	-0.122	0.158***	-0.190*
	(0.144)	(0.265)	(0.127)	(0.075)	(0.111)	(0.040)	(0.064)	(0.122)	(0.040)	(0.095)	(0.069)	(0.014)	(0.086)
Drift	0.000**	0.000	-0.000***	-0.000***	-0.000	0.000***	0.000***	-0.000	0.000*	-0.000***	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\text{Acousticness}_{t-1}$	**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 23: Vector Autoregression (VAR) Model Results

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

Bibliography

References

Aguiar, Luis, and Joel Waldfogel. 2018. "Quality Predictability and the Welfare Benefits from New Products: Evidence from the Digitization of Recorded Music." *Journal of Political Economy* 126, no. 2 (April): 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): 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* [in en]. Cambridge, MA, USA: MIT Press, October. ISBN: 978-0-262-01128-0.

Aridor, Guy, and Duarte Gonçalves. 2022. "Recommenders' originals: The welfare effects of the dual role of platforms as producers and recommender systems" [in en]. *International Journal of Industrial Organization* 83 (July): 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.

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. ISBN: 979-8-4007-0104-7, accessed January 14, 2024. https://doi.org/10.1145/3580507.3597677. https://doi.org/10.1145/3580507.3597677.

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.

Bergantiños, Gustavo, and Juan D. Moreno-Ternero. 2025. "Revenue Sharing at Music Streaming Platforms." *Management Science* 71, no. 10 (October): 8319–8335. ISSN: 0025-1909, accessed November 2, 2025. https://doi.org/10.1287/mnsc.2023.03830. https://pubsonline.informs.org/doi/abs/10.1287/mnsc.2023.03830.

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" [in en] (December). Accessed March 26, 2022. https://doi.org/10.48550/arXiv.1901.09851. https://arxiv.org/abs/1901.09851v2.

Chace, Zoe. 2011. "How Much Does It Cost To Make A Hit Song?" [In en]. *NPR* (June). 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* [in en]. Cambridge, MA, USA: MIT Press, February. ISBN: 978-0-262-03890-4.

Fleder, Daniel, and Kartik Hosanagar. 2009. "Blockbuster Culture's Next Rise or Fall: The Impact of Recommender Systems on Sales Diversity." *Management Science* 55, no. 5 (May): 697–712. ISSN: 0025-1909, accessed March 10, 2025. https://doi.org/10.1287/mnsc.1080.0974. https://pubsonline.informs.org/doi/10.1287/mnsc.1080.0974.

Fletcher, Amelia, Peter L Ormosi, and Rahul Savani. 2023. "Recommender Systems and Supplier Competition on Platforms*." *Journal of Competition Law & Economics* 19, no. 3 (September): 397–426. ISSN: 1744-6414, accessed March 10, 2025. https://doi.org/10.1093/joclec/nhad009. https://doi.org/10.1093/joclec/nhad009.

Fox, Jeremy T., Kyoo il Kim, Stephen P. Ryan, and Patrick Bajari. 2011. "A simple estimator for the distribution of random coefficients" [in en]. _Eprint: https://onlinelibrary.wiley.com/doi/pc *Quantitative Economics* 2 (3): 381–418. ISSN: 1759-7331, accessed March 29, 2026. https://doi.org/10.3982/QE49. https://onlinelibrary.wiley.com/doi/abs/10.3982/QE49.

Goeree, Michelle Sovinsky. 2008. "Limited Information and Advertising in the U.S. Personal Computer Industry" [in en]. _Eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECT *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.

Hosanagar, Kartik, Daniel M. Fleder, Dokyun Lee, and Andreas Buja. 2014. *Will the Global Village Fracture into Tribes: Recommender Systems and Their Effects on Consumers* [in en]. SSRN Scholarly Paper. Rochester, NY, April. Accessed March 10, 2025. https://doi.org/10.2139/ssrn.1321962. https://papers.ssrn.com/abstract=1321962.

Lancaster, Kelvin J. 1966. "A New Approach to Consumer Theory." *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" [in en]. *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 en]. In *Proceedings of the 12th ACM Conference on Recommender Systems,* 31–39. Vancouver British Columbia Canada: ACM, September. 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.

Mortimer, Julie Holland, Chris Nosko, and Alan Sorensen. 2012. "Supply responses to digital distribution: Recorded music and live performances." *Information Economics and Policy,* The Economics of Digital Media Markets, 24, no. 1 (March): 3–14. ISSN: 0167-6245, accessed September 28, 2024. https://doi.org/10.1016/j.infoecopol.2012.01.007. https://www.sciencedirect.com/science/article/pii/S016762451200008X.

Reimers, Imke, and Joel Waldfogel. 2023. *A Framework for Detection, Measurement, and Welfare Analysis of Platform Bias.* Working Paper, 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. 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* [in en-US], May. 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., C. Lanier Benkard, and Benjamin Van Roy. 2008. "Markov Perfect Industry Dynamics With Many Firms" [in en]. _Eprint: https://onlinelibrary.wiley.com/doi/pdf/1 *Econometrica* 76 (6): 1375–1411. ISSN: 1468-0262, accessed March 26, 2025. https://doi.org/10.3982/ECTA6158. https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA6158.