What drives demand for playlists on Spotify?

Max J. Pachali[1], Hannes Datta[1]

[1] Tilburg University

Stockholm School of Economics
May 10, 2022

Concerns about market power of digital platforms

- Today, online platforms play an important role in society and economy
- Platforms have started to dominate many industries
  - online retailing (e.g., Amazon)
  - search and search advertising (e.g., Google)
  - social media (e.g., Meta)
- Policy-makers around the world call for heightened transparency
  - U.S. Justice Department accuses Google of "illegally protecting its monopoly over search and search advertising" (20 Oct. 2020)[1]
  - U.S. Congressional Hearing: Amazon, Apple, Facebook and Google wield market power "to crush competitors and amass data, customers and sky-high profits" (30 Jul 2020)[2]
  - EU Observatory on the Online Platform Economy[3]

[1] "U.S. Accuses Google of Illegally Portecting Monopoly", 20 Oct. 2020, https://www.nytimes.com/2020/10/20/technology/google-antitrust.html

[2] "Amazon, Apple, Facebook and Google grilled on Capitol Hill over their market power", 30 Jul 2020, https://www.washingtonpost.com/technology/2020/07/29/apple-google-facebook-amazon-congress-hearing/

[3] https://ec.europa.eu/digital-single-market/en/policies/online-platforms

Platform power in the music streaming market

- Platforms have overtaken traditional revenue streams of content producers[4]
- In the digital age, users mostly consume music through listening to platform-curated playlists
- Concerns about the market power of music streaming services started after the rise of playlists on Spotify — the leading service[5]
- "Playlists are the top tool Spotify is currently employing to expand their platform empire"[6]
- Yet, not much is known about the drivers of users' playlist consumption on Spotify

[4] In 2019, 80% of content producer's revenues generated through online streaming.

[5] Market shares: Spotify: 36%, Apple Music: 18%, Amazon Music 13%, Tencent Music: 10%, Google/YouTube Music 5%, see https://musically.com/2019/12/09/report-spotify-has-36-market-share-of-music-streaming-subs/.

[6] Quote by Liz Pelly (https://tisch.nyu.edu/about/directory/clive-davis-institute/elizabeth-pelly) taken from https://watt.cashmusic.org/writing/thesecretlivesofplaylists, accessed October, 19, 2020.

Goal of our paper and data

- Understanding the drivers of playlist demand mirror power (im-)balances in the music streaming market
- If users mainly listen to a playlist because...
  - it contains **songs of popular artists**
    - $\rightarrow$ not Spotify would be powerful, but (major) record labels that have signed popular artists
  - of **its theme, mood, music topic, or overall acoustic style**
    - $\rightarrow$ playlist curators could exercise power through the decisions of which artists to include
  - it gets **featured in the app**
    - $\rightarrow$ Spotify would gain additional power, compared to other curators
- We empirically quantify the **main drivers of playlist choice** on Spotify in a structural demand model
  - unique daily data set of initially 1.2 million playlists (Oct. 2019 - March 2020, collected from *Chartmetric.com*)
  - complemented with daily information about which playlists get featured in the Spotify app (scraped from *everynoise.com/worldbrowser.cgi*)

Preview of Results

1. Persistent (time-invariant) preferences for Spotify playlists
   - → How much "better" does Spotify curate their playlists? If we set playlist preferences to the equivalent taste of non-Spotify playlists, we find that Spotify's market share in playlist consumption would drop by **15%**.
   - → Strong brands of Spotify playlists formed by professional editors who design a theme for every playlist
2. Responsiveness to featured playlists on Spotify's Search page
   - → Market share of Spotify playlists would drop by **6%** if the platform wouldn't highlight a very selective set of its playlists to users in the app
3. Valuation of the popularity of added songs and "This Is"-playlists of popular artists
   - → Due to the strong persistence of user preferences, Spotify playlists would only lose between **1%-2%** of market share if Universal (Sony, Warner) content was removed, and "This Is"-playlists were dropped from the platform.

Related literature

1. Effects of **digitization** on production, consumption, and consumer welfare (e.g., Datta et al. 2018, Aguiar and Waldfogel 2018)
   - → We study playlist consumption, rather than (self-directed) **song consumption**
2. Spotify's **playlist power**: Focusing on 6 exceptionally popular playlists, Aguiar and Waldfogel (2018) document Spotify's power to influence song and artist success
   - → We decompose the main drivers of user demand for playlists using a representative sample of **thousands** of popular playlists
   - → Our data contains daily information about which **playlists get featured** on Spotify's Search page, and we find that users are quite responsive to these "promotions" of playlists
   - → Our findings shed light on the **power imbalances** between major music producers and Spotify
3. Literature on **media power**: Simonov and Rao (2021) estimate persistent preferences for state-owned media outlets in Russia that are the main reason for media power and censored reporting
   - → In our setting, users' have strong tastes for platform-curated playlists that are the main driver of Spotify's power over music producers

Focus on the playlist ecosystem at Spotify

- Playlists are curated lists of songs that are typically centered around a particular theme e.g., by genre ("Deep House"), mood ("Relax"), or activity
- Mean (median) play duration of playlists equal 8h (4.6h)
- Millions of lists available, curated by Spotify, major and indie labels and users
- Playlist curators change compositions (i.e., add or remove tracks) based on perceived fit with a playlist's brand image
- Content producers generate revenues per stream of their tracks on any playlist ("streamshare")

The Spotify playlist landscape

1. **Spotify editorial playlists**
   - Like "Latin Pop Hits" (non-personalized), or "Deep House Relax" (semi-personalized)
   - Curated by a team of music experts and genre specialists from around the world hired by Spotify
   - Spotify owns about 14,961 editorial playlists and attracts 63% of all playlist followers
2. **Spotify algorithmic playlists**
   - Automatically created for each user by Spotify's algorithms
   - Users access these playlists in the "Made for You" section in the Spotify app (e.g., "Discover Weekly", "Release Radar", "On Repeat", "Your Daily Mixes", "Family Mix")
   - "Discover Weekly" and "Release Radar" are updated weekly on every Monday and every Friday, respectively
3. **Professionally curated playlists**
   - The three major music labels and many independent labels also operate own playlist brands[7]
   - "Tastemakers" — artists, radio stations, brands, or influencers that curate music
4. **User playlists**
   - Users can also share their playlists with one another and, in doing so, generate "followers" for their own libraries

[7] Sony (with Filtr), Universal (with Digstr), and Warner Music (with Topsify).

Data

Playlist data

- 1.2m playlists and daily data obtained from *Chartmetric.com* (October 2019-March 2020)
- Follower distribution across playlists is very skewed: the top 1% (12,153) of playlists add up to a follower share of 84% of the 1.2 million lists
- About 50% of all playlists tracked have 1 follower (i.e., they are not shared with other users)

Table: Playlist curators in full sample.

Playlist Curator	Follower share	Number of playlists	Mean followers
Spotify — Not personalized	0.38	10 777	93 132
Spotify — Semi-personalized	0.25	4184	145 047
Professional curators	0.12	69 460	4937
Other	0.10	352 940	743
Artists	0.06	86 090	1965
Major label (Sony)	0.03	4591	14 643
Independent labels	0.02	31 575	1824
Major label (Universal)	0.01	4676	7441
Major label (Warner)	0.01	3172	7502
Users	0.007	647 835	31

Note: Playlist curators shown in descending order of their follower share.

Data

Playlist data over time

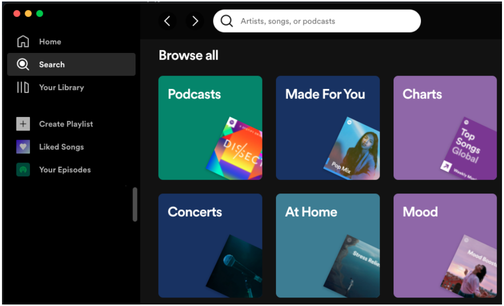
- Daily variation of playlist characteristics
  - → share of content owned by the three major labels (Universal, Sony, Warner)
  - → metric for the popularity of content on playlists: the number of tracks of a playlist added to other playlists at a given day
  - → number of tracks on playlists
  - → average time since release of tracks ("track age")
  - → average length of tracks (in minutes),
  - → nine numeric measures characterizing the acoustic attributes on playlists (danceability, energy, speechiness, acousticness, instrumentalness, liveness, valence, tempo, loudness)
- Daily variation of recommendation intensity at the platform level
  - → proxied by the daily refreshment[8] of Spotify's algorithmic playlists in our sample
- *Chartmetric.com*'s estimate of aggregate daily playlist listeners[9]

[8] Defined as the number of songs added to algorithmic playlists, divided by the number of algorithmic playlists in our sample.

[9] The company collects the number of listeners of artists added to the playlists. Daily listeners are then computed as the mean number of listeners of (observable) artists added to the playlist.

Data

Merge feature data to playlist data



- Spotify features a selected number of playlists on the Search page in their app
  → by country and categories (e.g., "mood", "dinner", "pop", "EDM")
- Collect daily data on featured playlists on Spotify from *Everynoise.com/worldbrowser.cgi* (October 2019 until March 2020)
  - data records a timestamp, category (e.g., "latin", "focus", "chill", "rock", etc.), the country name, and a list of playlist IDs featured in the category/country at the time of data retrieval
  - 6,739 playlists were featured during our period of observations in any of the sections and hence a very small fraction (0.56%) of all playlists
- Merge data on featured playlists with playlist data by IDs and dates

Data

Categorization of playlist ecosystem

- We use up to eight tag words (keywords) associated with playlists in the data to classify playlists into music genres using an association rule mining machine learning algorithm (e.g., Hastie et al., 2009)
- Our categorization builds on the notion that playlists with similar keywords belong to the same music genre and are hence considered as closer substitutes by users
- Start classification based on Spotify's pre-defined genres of featured playlists on the Search page (31 main genres)
- Repeatedly apply the Apriori algorithm (Agrawal et al., 1995) to identify the higher support item sets
- We mine in total 145 rules ("associations")

Data

Categorization of playlist ecosystem

- A playlist may be associated with multiple main genres, i.e. some keywords may create an association with rock, while other keywords create an association with romance
- Our final genre categorization takes associations to multiple main genres as separate clusters, i.e. rock & romance is a different category than rock
- Top 10 categories derived from the association rule learning classification:

Table: Top 10 Categories in the Data.

Category name	N	Followers (in m.)	Top 3 playlists
pop	58,962	378.37	Today's Top Hits, Global Top 50, Songs to Sing in the Car
latin	17,394	174.60	¡Viva Latino!, Baila Reggaeton, Rock en Español
pop+rb+student	73,123	130.09	All Out 00s, Songs to Sing in the Shower, I Love My '00s &B
hiphop	36,797	102.46	RapCaviar, Get Turnt, I Love My '90s Hip-Hop
rock	58,606	80.92	Rock Classics, 90s Rock Anthems, Legendary
chill+electronicdance+party	19,755	73.62	Beast Mode, Dance Party, Power Hour
folkacoustic+romance	22,617	56.59	Your Favorite Coffeehouse, Run Wild, Afternoon Acoustic
comedy+hiphop	47,747	51.44	This Is Drake, ThrowbackThursday, B.A.E.
chill+electronicdance+party+pop	12,475	43.79	mint, Motivation Mix, Dance Hits
rock+romance	7,291	37.34	All Out 80s, Have a Great Day!, All Out 70s

Note: Classification derived using association-rule mining algorithm.

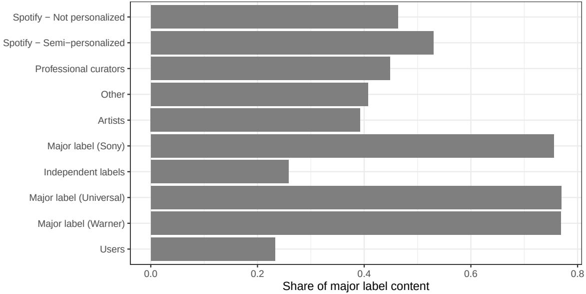
Data

Construction of estimation sample

- Our final estimation sample relies on 12,055 major playlists in 50 music categories covering 65% of the total followers of the 779,117 playlists we were able to allocate into genre categories[10]
  - → About 50% playlist were dropped that only had a single follower (no listener data available)
  - → Select the genre categories belonging the top 95% of cumulative follower market share of categories with at least 50 playlists included per category
  - → Select the 95% mostly followed playlists within considered categories
  - → These steps reduce the number of considered playlists to 30,746 in 81 genre categories covering 88% of followers of the 779,117 playlists
  - → Among the 30,746 playlists, *Chartmetric.com* only observes listener data for 12,055 playlists that define the estimation sample
  - → Interpolation needed for listeners (mostly short streaks of missing days)
- We do not rely on the follower data as a metric for daily demand as users can passively follow many playlists without actually listening to them on a given day

[10] We had to drop 436,183 playlists from our sample that did not contain any keyword.

Average share of major label (ML) content per curator type



- Labels push their own content (high ML share on major label lists, low ML share on independent label lists)
- Spotify lists have similar ML share as lists by professional curators, artists and "others"
- Even indie labels need major label coverage

Major label share, popularity, and listeners

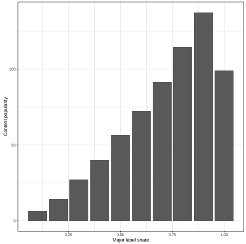


Figure: Content popularity and major label share

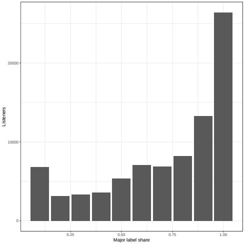


Figure: Listeners and major label share

- Adding major label content to playlists is associated with higher overall content popularity
- Playlists with a larger major label share seem to have more listeners

Over time variation at the playlist level

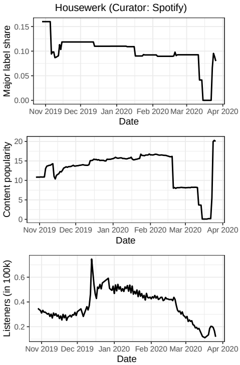


Figure: Housework (by Spotify)

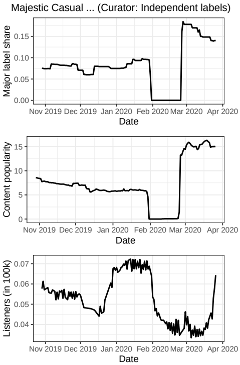
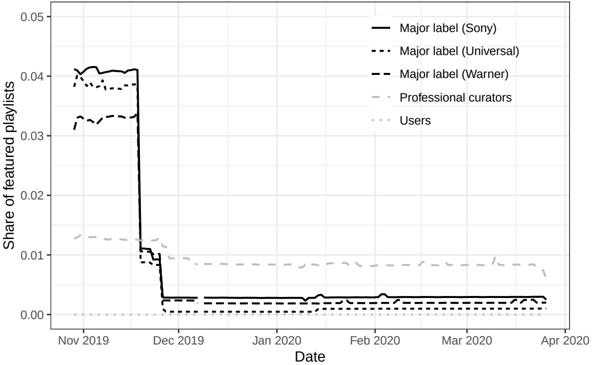


Figure: Majestic Casual (by an Indie Label)

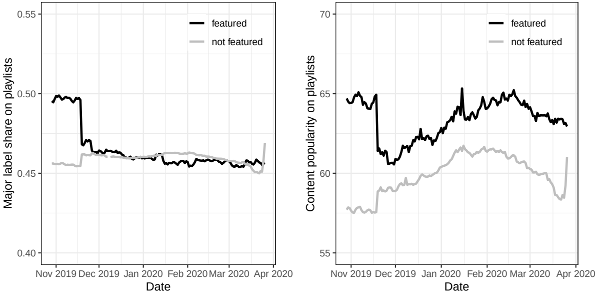
- The previous plots do not account for playlist-specific demand factors
- We rely on over time variation at the playlist level to identify demand effects
- Variations in major label share affect content popularity over time
- But also the outbreak of the Covid 19 pandemic reduced the popularity of house music tracks (e.g., Sim et al. 2022)

Variation in featured playlists across curator types



- The majority of playlists in our sample are never featured on Spotify's Search page (in at least one section in at least one country)
- By November 2019, Spotify decided to stop featuring major label playlist and substantially increased the support of its own playlists

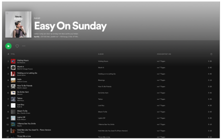
Major label share and listeners for featured vs. non-featured playlists

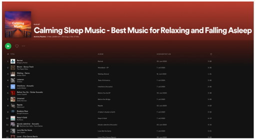


- Regime shift induced a substantial reduction of featured tracks released by any of the three major labels
  - Evidence for decreasing major label power on Spotify
  - Spotify aims to increase the market share of independent labels and pursues the "goal to incorporate new or smaller artists into a variety of playlists" (Spotify 2020)
- Featured playlists have a substantially higher number of daily listeners

Modeling playlist demand

- Playlists as "products" differentiated by acoustic attributes, content popularity, play duration, curator type, . . .
- Every playlist has its own theme or hypothesis and users have different preferences for these "brands"
- Users can choose among thousands of differentiated playlists and we model this decision as a discrete-choice
  - Daily choice to listen to a playlist on Spotify, or consume music at other music streaming platforms
  - "The market": Number of premium subscribers of seven leading music streaming platforms (Spotify, Deezer, Apple, Amazon, Tencent, Pandora, YouTube)[11]





[11] We neglect the impact of fees on platform choice (monthly fees are almost homogeneously set across platforms to $10; in 2020: Spotify $9.99, Apply Music $9.99, Deezer $9.99, Amazon Music $9.99 (None-Prime, without accounting for annual plan discounts), Pandora $9.99, YouTube Music $11.99.

Demand specification

- Every day t, every consumer i in the market chooses p from the $P_t + 1$ options that maximizes utility[12]

$$u_{ipt} = x_{pt}^T \Delta + \xi_{pt} + \zeta_{igt} + (1 - \sigma) \cdot \varepsilon_{ipt}, \tag{1}$$

- $x_{pt}^T \Delta$ is the mean utility component from playlist p at t (details on next slide)
- ξ_{pt} is an unobserved time-varying playlist popularity or quality term
- ζ_{igt} is a random utility term such that $\zeta_{igt} + (1 - \sigma) \cdot \varepsilon_{ipt}$ has an extreme value distribution

[12] Consumers can switch platforms on a daily basis in our model and additional fees from temporal double subscriptions are "sunk costs".

Empirical specification of consumer utility

$$\begin{aligned} u_{ipt} =& \gamma_p + \eta^{\text{feat}} n_{pt} + \beta pop_{pt} - \alpha c_{pt} + \omega r_t \\ & + \gamma_{M(t),C(p)} + \gamma_{D(t)} + \rho t + \kappa \mathbb{1}\{t \geq \text{Covid}\} + \phi^T x_{pt} + \zeta_{igt} + (1-\sigma) \cdot \varepsilon_{ipt}, \end{aligned} \tag{2}$$

- γ_p denotes users' time-invariant preferences for playlist p ("playlist brand equity")
- η^{feat} measures the effect of featuring playlists on Spotify's Search page; n_{pt} is the number of sections playlist p is featured in, aggregated over countries
- pop_{pt} proxies for "content popularity" using the number of tracks from the focal playlists p added to other playlists $-p$
- c_{pt} is the playlist duration (in hours) and α is the marginal utility of time
- ω measures the effect of personalized recommendations[13]
- $\gamma_{M(t),C(p)}$ represent month-category dummies, $\gamma_{D(t)}$ is a day-of-the-week dummy, ρ controls for a general trend ("platform growth") and κ accounts for the outbreak of the Covid 19 pandemic
- $\phi^T \boldsymbol{x}_{pt}$ controls for track age as well as acoustic attributes of p
- We partition playlists into $G = 50$ nests (obtained from the categorization using association rule mining) and σ is the nesting parameter

[13] Defined as the number of songs added to algorithmic playlists, divided by the number of algorithmic playlists in our sample.

Estimation

- Omitting the t index and setting $\delta_p := x_{pt}^T\Delta + \xi_{pt}$, the market share of playlist $p \in H_g$ is given as:

$$s_p = \frac{\exp\left[\delta_p/(1-\sigma)\right]}{\left(\sum_{h\in H_g}\exp\left[\delta_h/(1-\sigma)\right]\right)^{\sigma}\cdot\sum_g\left\{\left(\sum_{h\in H_g}\exp\left[\delta_h/(1-\sigma)\right]\right)^{1-\sigma}\right\}}, \quad (3)$$

- where H_g is the set of playlists within the nest g
- We invert market shares as proposed in Berry (1994) and estimate the model via OLS:

$$\begin{aligned}\ln s_{pt} - \ln s_{0t} =& \gamma_p + \eta^{\text{feat}} n_{pt} + \beta pop_{pt} + \sigma \ln s_{pt|gt} - \alpha c_{pt} + \omega r_t \\ &+ \gamma_{M(t),C(p)} + \gamma_{D(t)} + \rho t + \kappa \mathbb{1}\{t \geq \text{Covid}\} + \phi^T x_{pt} + \epsilon_{pt},\end{aligned} \quad (4)$$

- $s_{pt|gt}$ represents market share of playlist p in nest g and ϵ_{pt} corresponds to residual factors affecting demand for playlist p at day t

Endogeneity

- We use instruments to correct for the potential endogeneity in the timing of featuring playlists and within nest market share

- **Playlist featuring**: Daily number of sections on Spotify's Search page that very small playlists are featured in (belonging to the 10% of observations with the lowest followers in the same category)
- Intuition: Simultaneous coordination of the decision which playlists to feature within categories increases cost efficiency (relevance)[14], but very small lists unlikely affect demand of major lists (different markets: "mainstream" vs. "niche")

- **Within nest market share**: Sum of play duration (number of tracks $\times$ length in hours) of *competing* playlists at every day (BLP-type)
- Intuition: This instrument is exogeneous (i.e., unrelated to ϵ_{pt}) because the expansion of tracks on the platform rather reflects supply and not demand shocks, but negatively correlates[15] with $s_{pt|gt}$

[14] First-stage F-statistic: 63.76.

[15] First-stage F-statistic: 211.24.

Endogeneity

- Curators may change playlist compositions in response to daily shocks
- However, industry insiders (e.g., Heuvelings, 2020) describe that playlist compositions get updated at fixed intervals on Spotify — mostly at a weekly level
- The mean (median) average update frequency of playlists in our sample equals 11.8 (7.8) days

- The majority of curators thus do not adjust playlist compositions based on daily unobservables and the lion's share of variation in our data holds strategic decisions of curators constant
- We follow (Rossi 2018) who proposes to deal with potential endogenous variables by using data sampled at a much higher frequency than strategic decisions are made

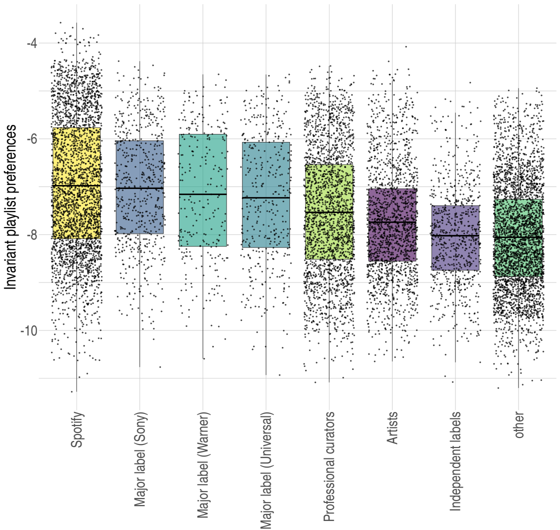
Nested logit regressions

Table: Playlist demand estimates.

	OLS		IV	
	(1)		(2)	
featured (# sections on search page)	0.001***	(0.000)	0.031***	(0.002)
content popularity	0.003***	(0.001)	0.045***	(0.009)
sigma	0.984***	(0.001)	0.719***	(0.022)
play duration (in h)	0.000***	(0.000)	−0.002***	(0.000)
track age	0.000 00**	(0.00000)	−0.000 03***	(0.00001)
trend	0.000***	(0.000)	−0.001***	(0.000)
covid outbreak	−0.016***	(0.000)	−0.001	(0.002)
recommendation intensity	0.003***	(0.000)	0.002***	(0.000)
Monday	0.000***	(0.000)	−0.005***	(0.001)
Tuesday	0.009***	(0.000)	0.003***	(0.000)
Wednesday	0.010***	(0.000)	0.004***	(0.000)
Thursday	0.006***	(0.000)	0.002***	(0.000)
Friday	0.003***	(0.000)	−0.002***	(0.000)
Saturday	0.007***	(0.000)	0.001**	(0.000)
ac - acousticness	0.035***	(0.011)	0.257**	(0.112)
ac - danceability	−0.027*	(0.016)	−0.110	(0.150)
ac - energy	−0.003	(0.019)	0.182	(0.170)
ac - instrumentalness	−0.004	(0.014)	−0.006	(0.084)
ac - liveness	−0.006	(0.014)	0.221*	(0.123)
ac - loudness	0.001	(0.001)	−0.009	(0.007)
ac - speechiness	−0.022	(0.023)	−0.320	(0.199)
ac - tempo	0.000	(0.000)	0.000	(0.001)
ac - valence	0.024**	(0.010)	0.282***	(0.104)
Observations	1,760,545		1,760,545	
R2	0.981		0.949	
Adj. R2	0.981		0.949	

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Robust standard errors in parentheses.

Estimates of playlist preferences



Counterfactuals

1. Compare the magnitudes of main factors of playlist demand: By how much would market share of Spotify-curated playlists drop ...
   (a) ... if they wouldn't be associated as such strong brands by users?
   (b) ... if they wouldn't be featured on Spotify's Search page?
   (c) ... if they wouldn't contain any (popular) Universal, Sony or Warner tracks and "This Is"-playlists of popular label artists would be removed from the platform (e.g., "This Is: Ed Sheeran")
2. Illustrate dependency of major label revenues on Spotify's editorial decisions
   (a) By how much would revenues of Universal, Sony and Warner drop would their content be removed from featured playlists on Spotify's Search page?

Impact of time-invariant playlist preferences

- Changes in playlist market share if users have substantially lower time-invariant preferences for Spotify playlists
- Reduce inferred preferences for Spotify playlists to the equivalent intercept of non-Spotify playlists

Table: Reducing users' time-invariant preferences for Spotify playlists.

Curator	Baseline	Counterfactual	Δ	95% CI
Outside good	64.69	72.40	7.72	[5.90, 9.29]
Spotify	30.94	16.38	−14.56	[-18.25, -10.50]
Professional curators	1.29	3.35	2.06	[1.34, 2.94]
Artists	1.20	3.06	1.86	[1.16, 2.67]
Major label (Sony)	0.65	1.74	1.09	[0.62, 1.73]
other	0.57	1.43	0.86	[0.54, 1.21]
Independent labels	0.27	0.57	0.30	[0.18, 0.48]
Major label (Universal)	0.23	0.63	0.40	[0.22, 0.67]
Major label (Warner)	0.17	0.43	0.26	[0.16, 0.43]

Note: This scenario reduces invariant user preferences for Spotify playlists to the inferred equivalent of non-Spotify curated playlists.

Impact of removing Spotify playlists from the Search page

- Changes in playlist market share if Spotify playlists don't get featured on the Search page

Table: Removing Spotify playlists from the Search Page.

Curator	Baseline	Counterfactual	Δ	95% CI
Outside good	64.69	68.31	3.62	[1.89, 6.26]
Spotify	30.94	24.71	−6.23	[-10.33, -3.42]
Professional curators	1.29	2.05	0.76	[0.41, 1.36]
Artists	1.20	1.87	0.66	[0.39, 1.11]
Major label (Sony)	0.65	1.05	0.40	[0.21, 0.72]
other	0.57	0.93	0.36	[0.19, 0.59]
Independent labels	0.27	0.43	0.17	[0.08, 0.30]
Major label (Universal)	0.23	0.39	0.17	[0.09, 0.34]
Major label (Warner)	0.17	0.26	0.09	[0.04, 0.19]

Note: This scenario simulates the effect if no Spotify playlist would be featured in the Search Page.

Impact of removing major label content

Empirical relation between major label share and content popularity

- Do Spotify playlists need popular major label tracks?
- Predict counterfactual content popularity of a playlist absent major label tracks, using splines:

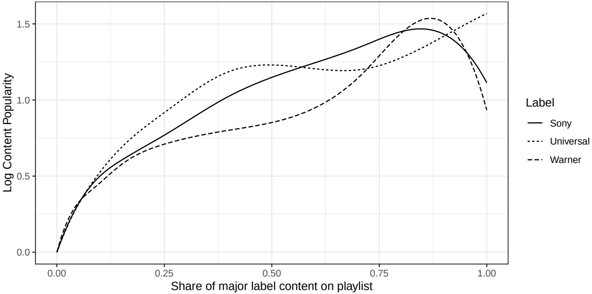


Figure: Fitted splines.

Impact of removing major label content

- Changes in playlist market share if Spotify playlists removed all tracks released by a major label, and if "This Is"-playlists of popular label artists were removed altogether (e.g., "This Is: Ed Sheeran")

Table: Removing major label content from Spotify playlists.

		No Universal			No Sony			No Warner		
Curator	base	counter	Δ	95% CI	counter	Δ	95% CI	counter	Δ	95% CI
Outside good	64.69	65.85	1.16	[0.71, 1.77]	65.52	0.83	[0.56, 1.08]	65.35	0.66	[0.46, 0.88]
Spotify	30.93	29.08	−1.85	[-2.90, -1.09]	29.64	−1.29	[-1.72, -0.79]	29.93	−1.00	[-1.42, -0.67]
Professional curators	1.29	1.50	0.20	[0.09, 0.36]	1.43	0.13	[0.06, 0.21]	1.39	0.10	[0.05, 0.17]
Artists	1.20	1.40	0.20	[0.08, 0.36]	1.34	0.13	[0.06, 0.24]	1.30	0.09	[0.05, 0.16]
Major label (Sony)	0.65	0.75	0.10	[0.05, 0.19]	0.73	0.08	[0.04, 0.13]	0.70	0.05	[0.02, 0.11]
other	0.57	0.66	0.09	[0.05, 0.15]	0.63	0.06	[0.03, 0.09]	0.61	0.04	[0.02, 0.07]
Independent labels	0.27	0.29	0.03	[0.01, 0.04]	0.28	0.02	[0.01, 0.03]	0.28	0.01	[0.01, 0.02]
Major label (Universal)	0.23	0.27	0.04	[0.02, 0.08]	0.25	0.02	[0.01, 0.04]	0.25	0.02	[0.01, 0.04]
Major label (Warner)	0.17	0.20	0.03	[0.01, 0.06]	0.19	0.02	[0.01, 0.03]	0.18	0.01	[0.01, 0.02]

Note: This scenario simulates the joint effect of (i) removal of Spotify's "This is"-playlists tailored to major label artists and (ii) removal of major label tracks from Spotify playlists.

Major label revenues and Spotify's editorial decisions

- Major label's $l \in \{\text{Universal, Sony, Warner}\}$ revenues from editorial playlists on Spotify can be expressed as follows:

$$\begin{aligned} R_l =& \mathit{streamshare}_l \times \mathit{payout}, \text{ with:} \\ \mathit{streamshare}_l =& \sum_{p \in P_{\text{-}0}} \frac{s_p\left(\delta_p\right)}{1-s_0} \times ms_p^l \\ \mathit{payout} =& R_{\text{Spotify}} \times z = (1-s_0) \times D \times f \times z \end{aligned} \tag{5}$$

- $P_{\text{-}0}$ is the set of playlists on Spotify (excluding the outside good), s_p is the market share of playlist p, ms_p^l is the share of tracks released by major label l on p, D is the market size, f is the subscription fee, and z is the share of royalties paid.

Major label revenues and Spotify's editorial decisions

- %-change of producers' revenues on Spotify in the three scenarios of no major label tracks on featured playlists

Table: Change of label revenues if Spotify removed major label content from its playlists.

Label	No Universal		No Sony		No Warner	
	Δ	95% CI	Δ	95% CI	Δ	95% CI
Universal	-88.15	[-92.28, -83.38]	3.45	[1.13, 6.18]	2.92	[0.46, 5.85]
Sony	6.46	[2.69, 11.73]	-82.05	[-88.84, -70.75]	3.49	[1.24, 6.75]
Warner	5.87	[1.86, 11.53]	4.21	[1.62, 7.64]	-91.36	[-94.47, -87.96]
Independent labels	49.85	[41.19, 62.17]	34.60	[28.42, 42.52]	27.30	[22.42, 33.10]

Summary of Results

- We study the main drivers of playlist demand on Spotify using a unique data set on daily listeners and characterizing attributes of about 12,000 playlists

- We find that users mainly chose Spotify playlists for time-invariant reasons, most likely because of their attractive themes
- Users are also quite responsive to featuring playlists on the platform's Search page → another part of Spotify's power that has not received much attention in the literature
- In relative terms, user demand is less responsive to fluctuations in the popularity of added tracks on playlists

- Robustness check: Consider order in which content is consumed (focus on top 10 tracks on a list only) — results are virtually the same

Discussion

(1) Revenues of major labels (and music producers more broadly) largely depend on Spotify's decisions about playlist compositions and featuring playlists on the Search page
  - everything equal, pitching tracks continues to be a good strategy
  - lack of featuring is worrisome → ads? regulator?

(2) In relative terms, Spotify is much less dependent on the availability of popular major label tracks on its playlists
  - playlist brand is key to Spotify's market success
  - labels can invest in broad catalogue of non-superstar content, but also need to "believe" such content matters

(3) Strategies to reduce Spotify's "playlist monopoly" and increase competition
  - curators should further invest in growing both the quantity and brand equity of their own playlist channels on Spotify; "playlist SEO"
  - at the same time, they need to find ways to affect streams outside of Spotify's editorial playlist ecosystem → influence algorithmic plays via ads?

Limitations

- Data
  - Large coverage of initially 1.2m playlists — but listener data absent or sparse for many
  - Aggregate data — individual-level and multi-platform data *at the country level* may capture choice more accurately
  - Only some algorithmic playlists included — studying these lists more may lead to strategies for mitigating platform power
- Model
  - Focus on short-term effects — future research may consider long-term effects (e.g., artists switching platforms, dynamics in playlist preferences)
  - Assumed equal consumption of tracks across playlists

Thank you for your attention!

SCAN ME

Business Model of Digital Music Platforms

Spotify as a two-sided market

- Spotify
  - Market demand depends not only on consumer preferences, but also on search costs
  - Platform lowers search costs and affects discovery (Aguiar and Waldfogel 2018), by means of playlists, recommendations, and curated artist profiles.
- Record labels
  - Major record labels account for 70% of global industry revenue (Universal 29.8%, Warner 22.6%, Sony 18.1%), and provide most of the highly popular content to Spotify
  - Independent record labels: highly fragmented (10,000+)
- Revenue model[16]
  - Consumers pay monthly subscription fee, and consume music via playlists (30% of consumption) or personal music libraries (70% of consumption)
  - Spotify collects revenue from ad-based and premium tiers; deducts operational expenses and pays out about 52% to labels, *relative to a label's number of streams* → labels compete for plays

[16] https://heroic.academy/artist-guide-spotify-playlist-royalties-verified-profiles/

Industry debates relationship between Spotify and the major labels

- Payout relative to plays on platform seems democratic and fair, but...
- Major labels own(ed) about 6-10% of Spotify shares, and major-label content is critical to the business model of Spotify ("can't do without Justin Bieber")[17]
- Consequently, major labels negotiated better deals with Spotify, compared to independent labels[18]
  - Higher-than-standard royalty rates (minimum per-stream payouts)
  - Control over share of advertising space, playlist addition, and playlist promotion
- Public debate on (fairness of) Spotify's business model
  - Shouldn't independent music artists earn the same, compared to major-label artists?
  - Shouldn't Spotify try to become label itself to reduce major-label power?
  - How powerful is Spotify in diverting attention (e.g., by introducing a new mood music category)?

[17] https://www.musicbusinessworldwide.com/one-reason-why-spotifys-deals-with-the-major-labels-rest-on-a-knife-edge/

[18] Sony contract leak in 2015, see https://www.theverge.com/2015/5/19/8621581/sony-music-spotify-contract.

Data

Categorization of playlist ecosystem

- We use up to eight tag words (keywords) associated with playlists in the data to classify playlists into music genres using an association rule mining machine learning algorithm (e.g., Hastie et al.,2009)
- Our categorization builds on the notion that playlists with similar keywords belong to the same music genre and are hence considered as closer substitutes by users
- Span the $(N \times P)$ matrix denoted as X with elements $x_{n,p} \in \{0, 1\}$, indicating whether a playlist n contains keyword p
- The goal of the algorithm is to find joint values of X that appear most frequently ("mode finding")
- We seek to find subsets of the integers $\mathcal{K} \subset \{1, \ldots, P\}$ such that

$$Pr\left[\bigcap_{k\in\mathcal{K}} (X_k = 1)\right] = Pr\left[\prod_{k\in\mathcal{K}} X_k = 1\right] \tag{6}$$

is large, where $\mathcal{K}$ is called an item set representing playlist keywords in our application

Data

Categorization of playlist ecosystem

- We seek all item sets $\mathcal{K}_l$ whose support exceeds the threshold $t = 0.1$:

$$\{\mathcal{K}_l | T(\mathcal{K}_l) > t\}, \text{ where}$$
$$T(\mathcal{K}_l) = \hat{P}r\left[\bigcap_{k \in \mathcal{K}_l} (X_k = 1)\right] = \frac{1}{N}\sum_{i=1}^{N}\prod_{k \in \mathcal{K}_l} x_{ik} \tag{7}$$

- Start classification based on Spotify's pre-defined genres of featured playlists on the Search page (31 main genres)
- Repeatedly apply the Apriori algorithm (Agrawal et al., 1995) to identify the higher support item sets (Equation 7) from the subset of playlists that contain the main genre as keyword, i.e. $\mathcal{K}_l^g(\tilde{X})$ with $\tilde{X} = \{X | \textit{main genre} \in \textit{keywords}\}$ [19]
- We then cast set of association rules surviving minimum confidence and lift:[20]

$$\mathcal{R}_g = \{A \rightarrow B | \mathcal{K}_l^g\}$$

- Stack total rules $\mathcal{R} = \{\mathcal{R}_g\}$

[19] By conditioning on a main genre at each iteration, we ensure that our classification is complete and consistent with Spotify's specification on the Search page.

[20] We set the confidence to 0.9 and the minimum lift of rules to 1.

Data

Categorization of playlist ecosystem

- We mine in total 145 rules for $A \rightarrow B,\ A \cup B = \mathcal{K}$, requiring that the consequent (B) is a main genre
- A playlist may be associated with multiple main genres, i.e. some keywords may create an association with rock, while other keywords create an association with romance
- Our final genre categorization takes associations to multiple main genres as separate clusters, i.e. rock & romance is a different category than rock
- Top 10 categories derived from the association rule learning classification:

Table: Top 10 Categories in the Data.

Category name	N	Followers (in m.)	Top 3 playlists
pop	58,962	378.37	Today's Top Hits, Global Top 50, Songs to Sing in the Car
latin	17,394	174.60	¡Viva Latino!, Baila Reggaeton, Rock en Español
pop+rb+student	73,123	130.09	All Out 00s, Songs to Sing in the Shower, I Love My '00s &B
hiphop	36,797	102.46	RapCaviar, Get Turnt, I Love My '90s Hip-Hop
rock	58,606	80.92	Rock Classics, 90s Rock Anthems, Legendary
chill+electronicdance+party	19,755	73.62	Beast Mode, Dance Party, Power Hour
folkacoustic+romance	22,617	56.59	Your Favorite Coffeehouse, Run Wild, Afternoon Acoustic
comedy+hiphop	47,747	51.44	This Is Drake, ThrowbackThursday, B.A.E.
chill+electronicdance+party+pop	12,475	43.79	mint, Motivation Mix, Dance Hits
rock+romance	7,291	37.34	All Out 80s, Have a Great Day!, All Out 70s

Note: Classification derived using association-rule mining algorithm.