How Curation Shapes Demand for Digital Information Goods: Estimating Playlist Elasticities for Music Streaming Services

Nils Wlömert
Institute of Retailing and Data Science, Vienna University of Economics and Business,
Welthandelsplatz 1, 1020 Vienna, Austria,
telephone: +43 1 31 336 4958
nils.wloemert@wu.ac.at

Dominik Papies
School of Business and Economics, University of Tuebingen,
Nauklerstr. 47, 72074 Tuebingen, Germany,
telephone +49 7071 29 78202
e-mail: dominik.papies@uni-tuebingen.de

Harald J. van Heerde
School of Marketing, UNSW Business School,
UNSW Sydney, NSW 2052 Australia,
telephone: +61 466 827 823
e-mail: h.vanheerde@unsw.edu.au

Acknowledgements: We are grateful for helpful comments from Stephan Seiler and seminar participants at the University of Maryland, the Texas Christian University and the European Quant Marketing Seminar (eQMS). We thank GfK and a music label that wishes to remain anonymous for granting access to the data.

ABSTRACT

Users of access-based services for digital information goods such as Spotify and Netflix are confronted with seemingly unlimited choice with thousands of movies and millions of songs. To help users navigate, platform operators offer curated selections. On music streaming services, for example, thousands of playlists, covering different genres, artists, and moods are available. This paper studies how the inclusion of a song on curated playlists affects this song's demand and how content owners (e.g., artists or labels) can strategically use these lists as a marketing tool. To this end, we collect information about the weekly curation of more than 60,000 playlists from a major streaming service which we relate to weekly streaming data for a sample of 54,000 songs over a period of 3 years, for a total of 7 million observations. Overall, we find a strong positive playlist effect with an elasticity of .20. Our analysis suggests stronger effects for playlists with more followers, featuring similar content from various artists, and with context-based curation. Furthermore, songs by less established artists and older songs benefit more from playlists. These playlist effects are stronger than the effects of traditional advertising, highlighting the vital role of curation in settings with abundant choice.

Streaming services like Apple Music or Spotify have assortments of unprecedented size that cover dozens of millions of songs. In fact, at the end of the first quarter of 2021, Spotify reported a catalogue size of more than 70 million tracks. Such a huge assortment of digital information goods such as music poses significant challenges for consumers because it makes the choice of which song to listen to very difficult. This has been referred to as the “tyranny of choice” (e.g., Mulligan 2015; Schwartz 2004a), and it is generally well established that large assortments often result in difficult choice tasks (Schwartz 2004b; Chernev, Böckenholt, and Goodman 2015; Scheibehenne, Greifeneder, and Todd 2010, Natan 2021).

Abundant assortments have become more prevalent due to the digitization of information goods and their online availability. Whereas in the analogous era, brick and mortar stores could only carry a very limited number of products (e.g., books, CDs, videos), the rise of digital information goods (e.g., MP3 files for music, digital books, streaming music, shows or movies) and the rise of online platform offering them means that there is practically no restriction on the shelf space and the number of products that retailers can offer. This has led to an explosion of assortments (Anderson 2004, Brynjolfsson, Hu, and Smith, 2010).

To help users navigate abundant choice, providers of digital goods have started to offer curated collections. Curated collections are defined as subsets of larger assortments that are grouped together based on one or more shared characteristics or benefits. In, e.g., the music streaming context, playlists have emerged as a tool for bundling music to support consumers in making choices from large assortments. By now there are about half a million playlists available to consumers on Spotify, with some playlists attracting more than 25m followers.

Curated collections are potentially important to help shape consumer demand. In access-based consumption models, content providers usually get paid based on the number of times their content is accessed, i.e., the remuneration for music artists in streaming services is based on how often their songs are streamed. This means that content providers (e.g., music artists) and their representatives (e.g., music labels) are highly interested in ways they can stimulate the usage of their content. At the same time, however, these parties (e.g., artists and music labels) can no longer attract customers with marketing instruments that are effective in other channels (e.g., pricing or promotions). The reason is that streaming services operate on a subscription base, and consumers do not pay for the consumption of individual songs, i.e., the marginal cost of a song for a consumer is zero. As pricing is not available as a marketing instrument, curated collections are a potentially powerful way to help drive demand. Realizing this potential, Universal (the world's largest music label) has dedicated teams creating playlists and pushing songs of their artists on influential playlists. Thus, playlists and their curation are emerging as an important new marketing tool.

Despite the potentially high commercial and user relevance of curated collections, the academic literature does not offer a comprehensive investigation how they play a role in shaping demand. One notable exception is Aguiar and Waldfogel (2021) who study the economic consequence of being listed on one of the 5 largest playlists on Spotify. In this paper, we go beyond that paper and offer two main contributions to the literature. First, we provide a first comprehensive analysis of the demand effects of playlists, an important emerging marketing tool relevant for ubiquitous digital information goods. By "comprehensive" we mean that we investigate a very large set of songs (54,000) across a large set of more than 61,000 playlists across a long time period (156 weeks or three years) for a census of streams in a large developed economy (Germany). This unprecedented dataset allows us to draw empirical generalizations on the demand elasticity of playlists and contrast it to more traditional elasticities that we also measure: advertising and radio airplay elasticities.

Second, our research sheds light on the boundary conditions or moderators of the demand elasticity of playlists. That is, drawing on analogies between the inclusion of songs on playlists and the inclusion of brands in retailer assortments, we document how the effect of playlist inclusion of a song on demand depends on song characteristics, playlist characteristics, and song-playlist characteristics. Leveraging quasi-experimental variation in the data and differencing out unobserved demand shocks, we find a playlist elasticity of around .20 that is much stronger than TV advertising (.008) and radio airplay (.017) elasticities. The playlist elasticity varies strongly in function of the song and playlist it pertains too. We show that the economic effect sizes of the inclusion of a song on a major playlist can be very substantial, generating potentially hundreds of thousands of dollars of additional streaming revenue.

In the remainder of this paper, we discuss the background literature, the rationale for the main effect and moderators of the playlist effect, the data, model-free- and model-based evidence, the revenue implications, concluded by a discussion.

LITERATURE

This research is related to at least three streams of literature. First, a substantial body of literature has established that choices become more difficult for consumers when more options are available (e.g., Iyengar and Lepper 2000; Natan 2021). Large assortments induce consumers to compare many options, resulting in very difficult choice tasks and feelings of unhappiness afterwards because consumers keep thinking about the options they could not choose (Schwartz 2004b). While there is no consensus about the average strength of a potential choice overload effect, Chernev et al. (2015) find that choice overload is a function of choice set size. Given that the assortment size in streaming services is much larger than the ones studied in previous research, it seems plausible that users of streaming services face a difficult choice task if they are not supported by appropriate tools to navigate the assortment. That is likely the reason why this situation has also been described as the “tyranny of choice” (e.g., Mulligan 2015; Shankleman

2021). Given the size of the choice task that consumers face in streaming services, it is likely that curation in the form of playlists are useful to consumers.

A second stream of research addresses the consumption and choice behavior of consumers in the context of digital music. An early study (Salganik, Dodds, and Watts 2006) demonstrates in the context of a music download store that consumers are susceptible to recommendations when they choose between different music titles. Datta, Knox, and Bronnenberg (2017) find that consumers indeed increase the variety of their music consumption once they join a streaming service such as Spotify. Most closely related to our research is a study by Aguiar and Waldfogel (2021) who study the role of playlists on the music streaming service Spotify. They analyze the 5 largest playlists on Spotify and focus on tracks that appear in the Spotify Top 200 charts. They find a strong positive main effect of a playlist listing on demand for that song.

A third stream of research that is relevant for our study deals with distribution decisions in retailing. That is, the inclusion of a song in a curated collection (e.g., playlist) has similarities to the inclusion of a brand in a curated retailer assortment. Research in retailing has studied the role of distribution, store size, shelf location (e.g., Ataman et al. 2008). In addition, researchers have assessed the respective advantages of structuring assortments by attributes (Lamberton and Diehl 2013; Rooderkerk and Lehmann 2020). We will draw on this research in the following section as a unifying basis when we describe the conceptual framework.

THEORY

Main Effect of Playlist Inclusion on Demand

Digital platforms, e.g., in the context of music streaming, offer millions of options (i.e., songs) to consumers, making choices for consumers difficult. Playlists are one potential tool that can make it easier for consumers to navigate these millions of options. If playlists are successful in helping consumers make choices, being included on a playlist should be beneficial for the included tracks and should lead to an increase in demand due to several reasons. First, many consumers subscribe to playlists to make their lives easier and to listen to whatever new (or existing) content is offered by the playlist. Thus, when a song is included on a playlist and consumers listen to it in an automatic mode (e.g., shuffle play), the song will obtain automatically obtain streams it would not have received when it had not been included.

Second, being included on a playlist leads to an increased visibility and raises awareness such that it becomes easier for artists and their music to be discovered by consumers. In this way, playlist serve as an advertising and sampling tool. Because artists cannot directly advertise to consumers within music streaming platforms or attract demand by running price promotions (because all songs are available for zero marginal cost), playlists can be viewed as the only remaining promotional tool that can be used to steer demand. Third, we can view playlists through the lens of distribution management. In a traditional, brick-and-mortar retail distribution context, the key goal of a brand's distribution management is to make a brand available to consumers. In the playlist context, the artist's goal is to make songs available and visible to consumers. In a brick-and-mortar distribution context, this implies a curation decision by retailers to include a given brand into the assortment. In the context of playlists, this curation decision is the decision to list a song on a given playlist. All these arguments jointly suggest that being included on a playlist is beneficial for a song's demand. This is also supported by empirical evidence in Aguiar and Waldfogel (2021) who analyze the largest playlists on Spotify and identify a substantial increase in streams for a song in response to its inclusion on one of these top playlists.

Returning to the analogy between playlists and brick-and-mortar retail distribution, Table 1 explains how we see the connections, which is also the basis for identifying moderators, discussed next.

Moderators of the Playlist Effect

We expect that the effect of playlist inclusion on streams will not be equally strong across all songs, playlists, and song-playlist combinations, but the effect will depend on a set of moderators. To inform our choice of moderators, we again follow the analogy between playlists and the traditional brick-and-mortar retailing context as a unifying framework.

Table 1

Analogy between playlists and brick-and-mortar retail distribution

Concept	***Distribution context***	***Playlist context***
Goal	Make brand available to consumers	Make song available to consumers
Curation decision	Brand included in retailer assortment	Song listed on playlist
Focal metric	**Distribution elasticity**: % change in sales to 1% increase in distribution	**Playlist elasticity**: % change in number of streams due to inclusion in 1% more playlists
LT Elasticity (average)	.61 for existing brands (Ataman et al. 2010) .76 for new brands (Ataman et al. 2008)	THIS STUDY
Moderator of distribution elasticity	Many possible moderators, see Ailawadi and Farris (2020)	THIS STUDY

Distribution weight. A key variable in distribution is the size of the stores that a brand is featured in. This is usually measured with the all commodity values (ACV; Ataman, Mela, and van Heerde. 2008) or store traffic, leading to weighed distribution measures where the weight is proportional to the size of the store. In the playlist context, the analogy is the number of followers of a given playlist (*playlist popularity*). In line with the distribution analogy, we also expect the effect of playlist inclusions to be stronger for larger, more popular playlists.

Choice set size. In retailing, a key consideration concerns the size of the assortment, i.e., the number of different options in a store's assortment. The playlist equivalent would the *length of the playlist*, i.e., the number of songs featured on a playlist. Here, we expect a negative impact on the playlist effect, i.e., the beneficial effect of being listed on a playlist becomes weaker the more songs a playlist features. The reason is that very long playlists again create a choice overload problem of their own, and attention and streams have to be divided among many tracks (e.g., Chernev et al. 2015).

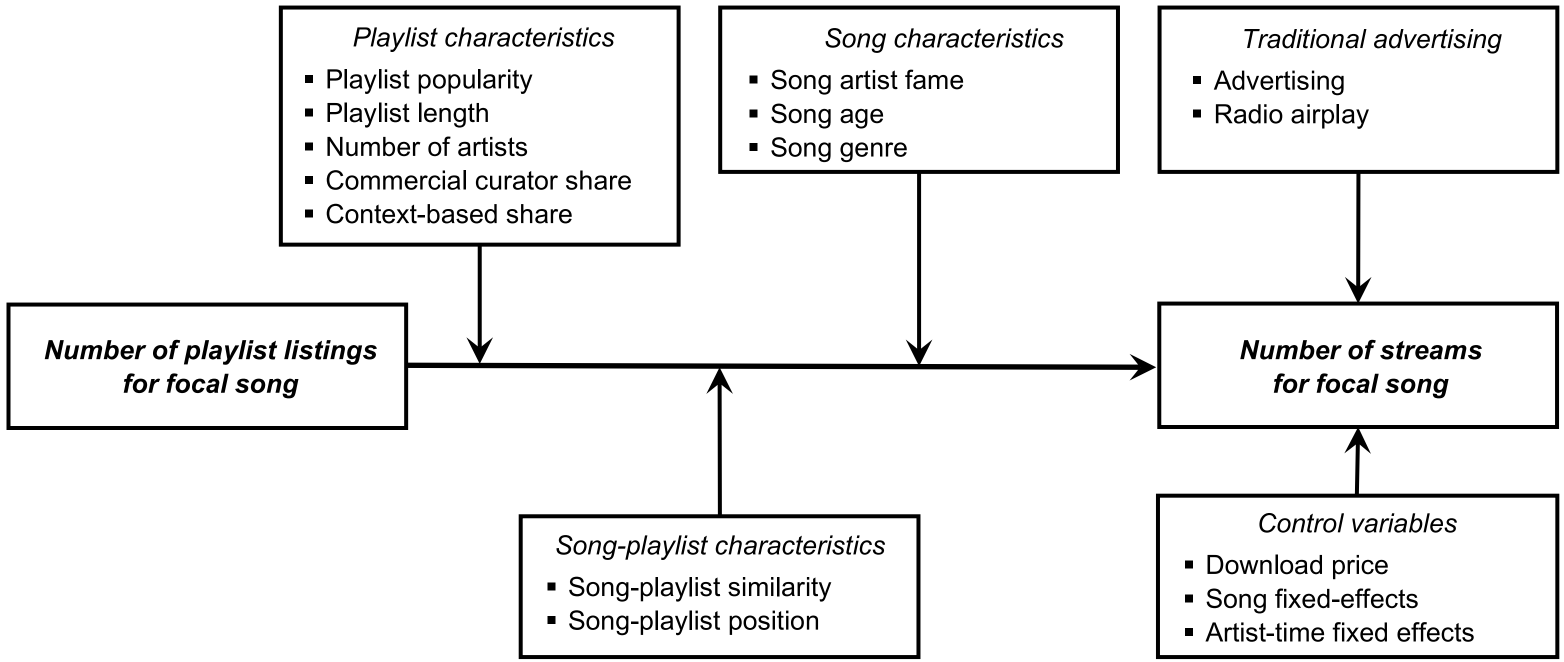
Brand range. A key aspect of large product assortments is that consumers are not fully informed about all products (Hendricks and Sorensen 2009) and it is important for artists and their products to be discovered. It is likely to be easier for artists to be discovered when they are listed on playlists that feature a variety of different artists because each of the co-listed artists may draw attention to the playlist and the chances of being exposed to new audiences who are not familiar with the artist yet increases. This is akin to promotional or advertising spillover (e.g., Sahni 2016), in which advertising for a focal brand spills over to a competing brand. We therefore expect that a given song will benefit relatively more from a playlist that features a larger number of artists.

Assortment control. In brick-and-mortar settings, retailers typically create and curate their assortments, although they may sometimes leave some of this responsibility to manufacturers acting as category captains. A unique characteristic of playlists is that playlists can be created and curated by anybody, i.e., by other consumers, music labels, experts and opinion leaders, or the streaming service itself. We do not have clear theoretical expectations of which types of playlists are more effective. On the one hand, non-commercial playlists by independent curators may be perceived as more trustworthy because the curators do not have vested interests (Erdem and Swait, 2004, Colicev et al. 2018, Goh, Heng, and Lin, 2013). Commercial curators, on the other hand, may have more and better resources available to identify music that matches the preferences of a playlist's set of followers.

Choice set organization. In traditional retailing, assortments can be organized by attributes (e.g., in clothing, this would be pants, shirts, skirts) or by benefits (e.g., clothing for work, or for weekends), and a coherent organization along these dimensions can mitigate potential negative effects of too much choice (Lamberton and Diehl 2013, p. 393).

Figure 1

Conceptual framework



In the context of playlists, a classical attribute-based organization would rely on artist names or genres. A more recent development has been the proliferation of playlists that focus on activities or contexts, e.g., a playlist for running, cleaning, or songs to play in the car. These context-based playlists usually assemble different artists from different genres and vintages, and they can be viewed as benefit-based choice set organization. We expect that an inclusion on benefit-based (i.e., context-based) is more beneficial for songs compared to non-context-based playlists. The reason is that it is likely that a context-based playlist familiarizes listeners with a song who would otherwise not have come across this song because, e.g., the song is outside the typical range of genres that the listener would typically choose.

Similarity to choice set. Another consideration a retailer faces pertains to the composition of brands in a category, i.e., whether the retailer should stock a set of similar brands with homogenous attributes, or a set of more heterogeneous products. Likewise, playlists can be composed of similar or dissimilar songs. On the one hand, being placed on playlists with similar songs means that artists can tap into a segment of consumers who are likely to prefer this type of music, e.g., a rap song will more likely meet interested listener on a rap playlist compared to, say, on a country music playlist. On the other hand, a playlist with a heterogeneous assortment may attract consumers who are not yet familiar with a given song, which implies that a song may be more likely to find new listeners on a heterogeneous playlist. For example, Oestreicher-Singer and Sundararajan (2012) find that in a recommendation network of book titles, recommendations have a stronger positive effect on the demand for recommended books if this book is from a different category than the focal book. In sum, we do not have clear theoretical expectations for the direction of this variable's moderating effect.

Position in choice set. Once a store has decided to list a brand in a given category, it has to decide where in a shelf space it will place the product, e.g., whether it will place the product in an easily accessible location that most store visitors will see (e.g., just below eye level), or in a less attractive position, e.g., towards the bottom of the shelf. Similarly, for playlists, tracks can be placed at the top of the playlist or at its end, and we expect the effect of playlist inclusion to be stronger for top placements.

Brand strength. In traditional retailing, the distribution elasticity likely varies depending on the strength of the focal brand. Similarly, in the context of music streaming, the effects of playlists on streams may depend on how well-known the focal artist is. We expect that less well-known artists can particularly benefit from being included on playlists because it is more likely that consumers will be aware of more well-known artists even without playlists. In other words, there is less (more) to discover through playlists when an artist is more (less) well-known.

Brand age. Brands differ with respect to their respective stage of the product life cycle, e.g., some brands are well established and matured, whereas other have just entered the market. This will likely affect the strength of the distribution effect. Similarly, in the case of playlists, older songs may experience different effects from being included on playlists compared to new releases. Specifically, we expect that older songs will benefit more from being included on playlists, and the conceptual arguments are similar as in the case of brand strength. The logic is that newly-released songs are more likely to be top-of-mind for consumers even without being placed on playlists, and this reduces the need to make them findable via playlists. In line with this expectation, Zhang (2018) finds that the increased level of content sharing after the removal of digital rights management (DRM) restrictions on music download services disproportionately benefitted older releases for very similar reasons. Therefore, we expect the playlist effect to be particularly pronounced for older songs.

Category. Lastly, distribution effects are likely heterogeneous across product categories, and similarly, we expect the playlist effect to be heterogeneous across music genres such as Rock, Pop, HipHop, and Country. We do not formulate expectations.

In Table 2, we summarize the constructs that we use in this research and their respective definition in the retailing and in the playlist context.

Table 2

Moderating factors of distribution/playlist effectiveness

Construct	*Distribution context*	*Playlist context*	*Exp. effect on effectiveness*
Distribution weight	All Commodity Value, Store traffic	Playlist followers	+
Choice set size	Assortment size	Playlist length	–
Brand range	Number of brands	Number of artists on playlist	+
Assortment control	Retailer (vs consumer)	Commercial (vs user)	+/–
Choice set organization	By usage context (vs by attribute similarity)	By usage context (vs by music genre)	+/–
Similarity to choice set	Similarity to other brands on shelf	Song-playlist similarity	+/–
Position in choice set	Shelf location (lower)	Playlist position (lower)	–
Brand strength	Brand strength	Artist fame	–
Brand age	Brand age	Song age	+
Category	Product category	Music genre	n.a.

DATA

Sample

We work with the census of all songs that were available in the German streaming market between 2017-2019 (3 years). For all songs, we have all weekly streams across all major streaming services as well as paid downloads from all major download stores. To avoid focusing on songs that were hardly ever played, we consider all songs that generated at least 1,000 streams in total across the entire observation period (this corresponds to approx. $ 6.00 in revenue, that generated at least one stream per week after release and one download in total, that have at least one public playlist listing per week, and that we observe for at least 13 weeks in total.

We complement this data set with data on initially roughly 500,000 playlists that are available to German consumers on the major streaming service in the market (i.e., Spotify).[1] The

[1]We note that our playlist data comes from Spotify and the streaming data comprises streams from all major streaming services (incl. Spotify). This is because GfK only reports streaming data aggregated across all streaming services. However, Spotify is the dominant service in the market (Mulligan, 2019) and accounts for the majority of

set of playlists covers a diverse set of content selections, ranging from major playlists with millions of followers (e.g., Today's Top Hits – 25m. followers, Global Top 50 – 14.6m. followers, RapCaviar – 8m. followers) to smaller niche playlists. From these playlists, we focus our analysis on those 61,000 playlists that have at least 1,000 followers. In addition, we collect information on traditional marketing activities, covering data on weekly TV advertising budget per song and weekly radio plays for each song, based on the Top300 radio charts.

To qualify for our sample, a song needs to be listed on at least one of the 61,000 playlists per week. The reason for this selection criterion is that we can only compute playlist characteristics for songs that are listed on playlists. This leads to a final sample of 54,020 songs from 13,390 artists across 156 weeks and a total of approximately 7.3m. song-week observations. This sample covers more than 130 *billion* streams during the observation period, which is slightly more than 60% of the entire German streaming market.

Measurement

In this section, we explain the measurements of our model variables. We report descriptive statistics for these variables in Table 3.

Dependent variable. The focal outcome variable is the *number of streams* for a given song i in week t. This variable is measured by GfK across all major streaming services in the market. On average, a song receives approx. 17.9k streams per week. An inspection of the percentiles suggests that the distribution is highly skewed, which is typical in entertainment media markets, where a small number of superstars often accounts for a disproportionate share of demand (Hendricks and Sorensen, 2009). As we will describe below in more detail, we require the number of *paid downloads* in music download stores (e.g., Amazon mp3) for a given song in a given

---

streams in our data. In addition, we find a strong correlation between the streams on Spotify and the streams across all services (r = .971, p < .001, n = 22,755; n songs = 1,825; n weeks (mean) = 9.88) for a subset of songs for which we could obtain streaming data for Spotify in Germany from a public source (*https://spotifycharts.com*).

week for identification purposes. These data come from GfK and cover all major download stores in the German market.

Key independent variable. The focal construct of interest is the *number of playlists* a song is listed on in a given week. On average, a song is listed on 53 playlists per week and this distribution is fairly skewed, similar to the distribution of streams, i.e., a small number of songs receives a comparatively large number of playlist listings. The number of playlist listings not only varies cross-sectionally (between songs), but also over time, as illustrated in Figure 2 for a set of example songs. The figure also shows that there is a strong correspondence between the number of streams and playlist listings. The main estimation challenge of our analysis will be to identify the causal relationship of the number of playlist listings on the number of streams.

Moderator variables. To test whether the effect of playlist listings on streams varies predictably according to the rationales presented above, we measure moderator variables in three groups according to our research framework, i.e., playlist characteristics, song characteristics, and song-playlist characteristics. With regard to the playlist characteristics, we measure the *popularity of the playlists* a song is listed on in a given week, using the audience size, i.e., the number of followers that each playlist has. Because songs are mostly listed on multiple playlists in a given week, we then take the average number of playlists followers per week across the subset of playlists that a song is listed on.

We measure *playlist length* as the average number of songs included on the playlists that song *i* is listed on in week *t* and the average *number of artists* featured on these playlists is computed accordingly. To analyze the impact of the curator type, we measure the *playlist commercial share*. A playlist can be curated by commercial curators (e.g., a music label or a streaming service itself), or by independent curators (e.g., music connoisseurs). For each playlist, we create an indicator variable that measures whether a playlist has a commercial curator or not. Then, for song *i* in week *t*, we calculate the share of playlists with commercial curators.

To investigate if the playlist effect is contingent on the type of content curation, we measure the *context-based share* of playlists for song *i* in week *t*. A context-based playlist is curated according to a specific context (e.g., workout, driving), whereas attribute-based curation is based on content-based criteria, such as genres or artist names. To classify a playlist as a context-based playlist, we create a dictionary of keywords based on the playlist titles. If a playlist title contains any of the identified context-related keywords, we classify it as a context-based playlist. Based on this classification, we then calculate the share of playlists with context-based curation for song *i* in week *t*. We report details regarding this classification in Web Appendix A.

With regard to the song characteristics, we measure *song artist fame* using all previous chart placements of the artist that is associated with song *i* as the sum over the inverse of weekly chart positions since 1974 in the official German Top 200 music charts. *Song age* is measured as the number of weeks since a song was released. To measure *song genre*, we use the genre information provided by GfK and aggregate this information to reflect 9 main genres.

With regard to the song-playlist characteristics, we base our measurement of *song-playlist similarity* on audio features provided by Spotify that describe the music style of each song across eight dimensions (https://developer.spotify.com/documentation/web-api/reference/#endpoint-get-audio-features). We first calculate for each feature the mean across all songs that are listed on a given playlist in week *t*, which gives a vector of eight mean values per playlist and week. Then, for a given song in a given week, we compute the cosign-similarity between a song's audio features and the playlist's vector of song features. We then average the cosign-similarity across all playlists that a given song is listed on. The *song-playlist position* is measured as the average position of a song *i* in week *t* across all playlists. To account for varying playlists lengths, we standardize the number of songs per playlists before computing this variable.

Figure 2

Number of weekly streams, playlists, radio plays, and marketing expenditures for eight example songs

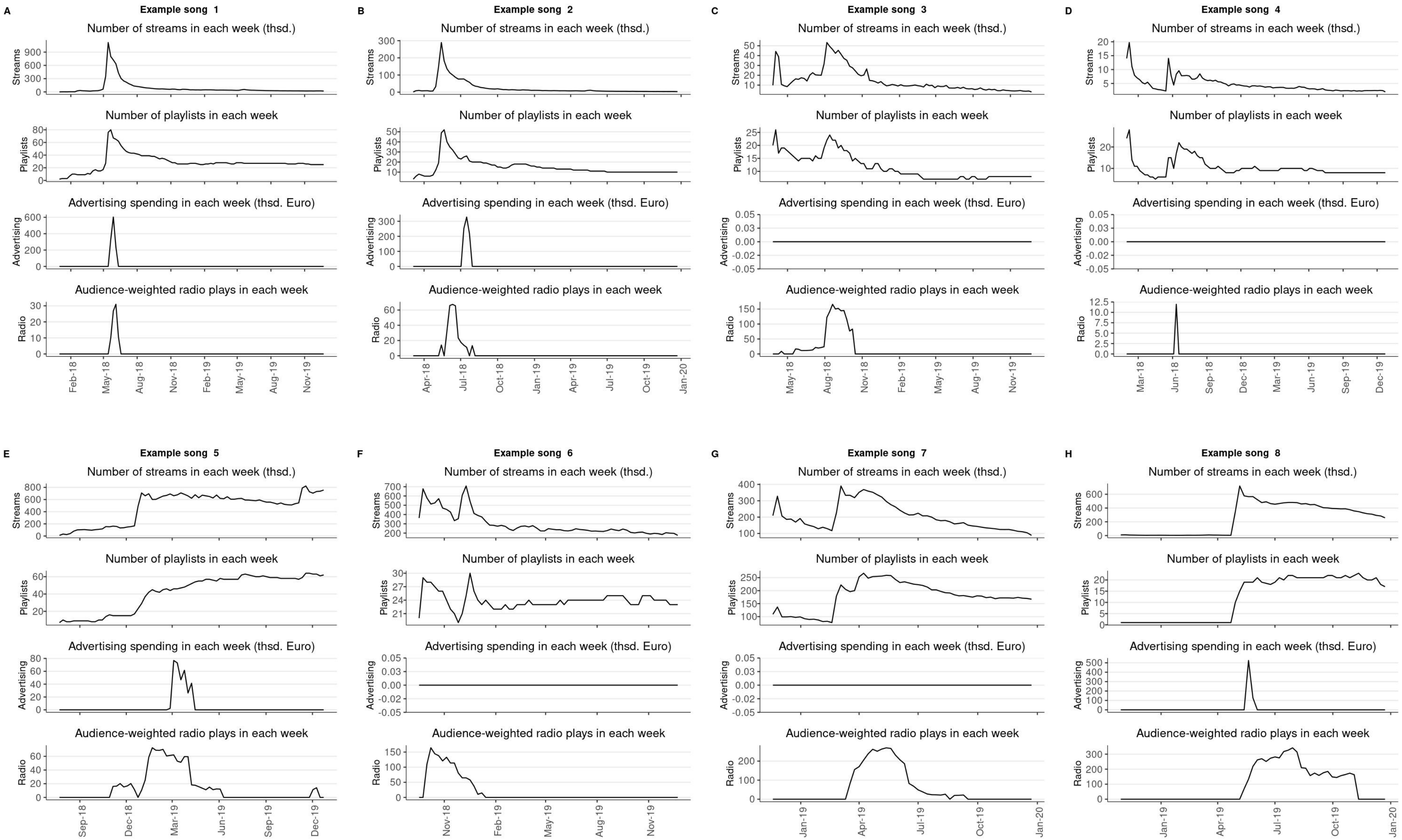


Table 3

Variable overview

Variable	*Definition*	*Role*	*Mean*	*SD*	*5%*	*25%*	*50%*	*75%*	*95%*
Number of streams	Number of streams	Numerator of DV	17,882.41	86,921.18	394.00	1424.00	3759.00	11,087.00	65,936.00
Number of downloads	Number of downloads	Denominator of DV	10.58	106.21	0.00	0.00	2.00	5.00	32.00
Number of playlists	Number of playlist listings	Focal IV	52.85	114.08	2.00	6.00	16.00	44.00	229.00
Playlist popularity	Average number of playlist followers	Moderator	20,179.48	35,135.16	2,244.60	5,825.50	10,466.36	20,216.08	69,848.99
Playlist commercial share	Share of playlists curated by commercial curators	Moderator	0.09	0.14	0.00	0.00	0.05	0.12	0.38
Playlist context share	Share of context-based playlists	Moderator	0.11	0.13	0.00	0.00	0.09	0.17	0.33
Playlist length	Avg. number of songs on playlist	Moderator	498.84	314.04	123.67	304.00	447.24	616.14	1,045.33
Playlist number of artists	Avg. number of artists on playlist	Moderator	224.08	137.07	49.00	132.78	205.50	288.33	458.61
Song-playlist similarity	Similarity of song to co-listed songs on playlists (cosine similarity based on 8 ‘audio features’)	Moderator	0.48	0.21	0.09	0.36	0.53	0.64	0.75
Song-playlist position	Relative position of song on playlist	Moderator	0.52	0.13	0.33	0.45	0.51	0.59	0.74
Song age	Number of weeks since a song was released	Moderator	608.41	653.77	23.00	120.00	363.00	880.00	1976.00
Song artist fame	Previous chart placements of artist (sum over the inverse of weekly chart positions since 1974)	Moderator	4.34	12.75	0.00	0.00	0.00	1.40	28.94
Advertising	TV advertising expenditures (weekly EUR budget)	IV	47.99	4,863.01	0.00	0.00	0.00	0.00	0.00
Radio plays	Number of weekly radio plays based on Top300 radio charts (audience-weighted weekly plays)	IV	0.92	24.66	0.00	0.00	0.00	0.00	0.00
Download price	The average price of a song	IV	1.25	0.38	0.99	1.29	1.29	1.29	1.29

Notes. N = 7,275,833. DV = dependent variable; IV =independent variable; Advertising expenditure and radio airplay is highly concentrated, i.e., only 326 songs out of the sample of 54,020 songs (.6%) receive TV advertising, and 2,687 songs (4.9%) enter the Top300 radio charts; even for songs that receive advertising and/or radio airplay, we only observe positive values for a fraction of weeks, which explains why the values for the 95th percentile of the respective variables is .00.

TV Advertising and *radio airplay*. We measure TV advertising as the weekly media spend (Euro amounts) of all TV spots featuring the artist that is associated with track *i* in week *t.* We measure radio airplay for track *i* in week *t* as the number of audience-weighted radio plays from the top 300 radio charts. Figure 2 shows the developments of advertising expenditure and radio airplay over time for selected example songs. It can be seen that while there is also a co-movement between the traditional advertising variables and the number of streams, the correlation between the number of playlists and streams appears stronger. It is also interesting to note the high concentration in the advertising and radio play variables (see Table 3). While all of the 13,390 artists in our sample receive playlist listings over the observation period, only 213 artists (1.6%) receive TV advertising expenditures and 1,282 artists (9.6%) receive radio airplay.

Control variables. We adjust for the *download price* by considering the average market price across all download stores in Germany in week *t* for track *i*. In addition, we adjust for unobserved song-specific effects by including fixed effects on the song-level. Furthermore, we adjust for any time-varying artist-specific popularity shocks by including fixed effects on the artist-week level. We present details regarding our model specification below.

ESTIMATION AND IDENTIFICATION

Our primary interest lies in the effect of a focal song's playlist listings on the focal song's demand, measured as the number of streams. In the absence of clear exogenous shocks to the number of playlists that a song is listed on, we face identification challenges that we discuss and tackle below. The main challenge arises because even with a rigorous fixed-effects structure it is possible that there are other, remaining unobserved time-varying factors at the song level that cannot be accommodated with fixed effects. This may happen if only the focal song experiences, e.g., an increase in popularity, but not the other songs of the same artist. One likely example is that a song is listed on a playlist because it has received many streams, or, more generally, that songs that are on an upward trajectory in terms of the number of streams get listed on playlists.

Both concerns share the underlying problem that a change in song-level streams may precede listings (reverse causality) or coincide with listings (simultaneity). These concerns arise because songs are not randomly placed on playlists but it is likely that curators pick songs that they believe are attractive for playlist followers. Hence, a positive correlation between playlist listings and streams is not necessarily causal because it may merely reflect the curators' ability to select songs that will be more successful in the future (Eliashberg and Shugan 1997). To tackle these concerns, we utilize the richness of the data to infer a causal effect of playlist listings on demand through different approaches that rely on different identifying assumption.

As a first step, we present model-free evidence of a causal link between playlist listings and the number of streams that a song receives. As a second step, we isolate playlist listing events with a pre- and post-period without playlist listings on the song level to construct quasi-experimental settings, which allows us to estimate the effect of playlist listings on streams using a difference-in-difference estimator. This step allows us to establish a measure for the overall effect size. In a third step, we specify an econometric model that addresses the concerns we discuss above, but that also allows us to assess the role of a broad set of moderators and potential dynamic effects. This econometric model differences out demand shocks observed in a different measure for demand: paid downloads.

Model-free evidence

To provide first evidence for a causal link between a song's playlist listings and the number of streams that this song receives, we revisit Figure 2, which shows the developments of our dependent variable (i.e., the number of streams) and our key independent variable (i.e., the number of playlist listings) over time for 8 example songs. The trajectories of both variables show a clear co-movement, providing support for the notion that there is a positive relationship between these variables. However, this pattern does not provide sufficient cues regarding the nature of this relationship, i.e., to what extent the change in streams is caused by a change in playlist listings. The co-movement could, for example, be driven by unobserved factors (e.g., song-specific popularity shocks) that affect both the number of streams and the number of playlist listings. Or an increase in streams may precede an increase in the number of playlist listings if playlist curators select songs that are already on a positive trajectory.

To provide more compelling evidence for a causal effect of playlist listings on the number of streams, we thus turn to Figure 3, which shows the number of streams per week for 6 example songs in our data set. There are many similar patterns for other songs in our data and we selected these examples to reflect the heterogeneity of our data in terms of song and playlist types. Each of the highlighted songs received a listing on one of the 61k playlists in our sample over our observation period. The grey area indicates the weeks in which this particular song was included on this particular playlist.

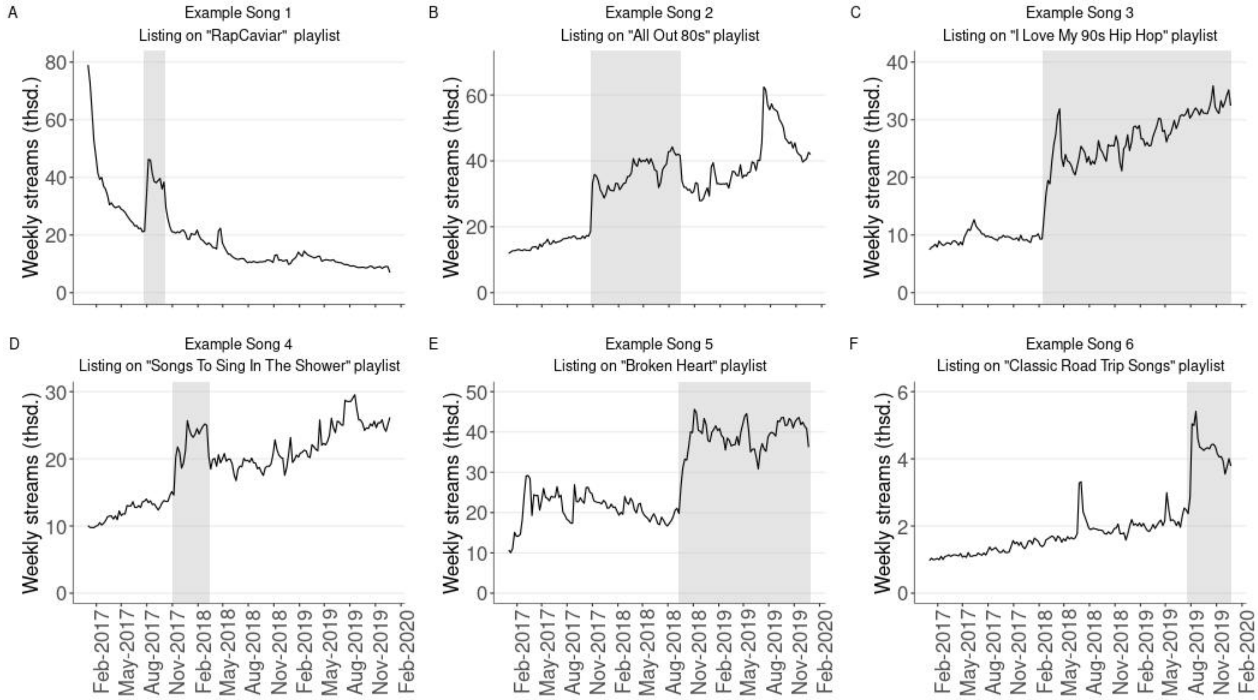
Figure 3, Panels A-C show examples of songs that got listed on playlists featuring songs according to a specific genre and/or based on vintage. Panel A shows a newly released song getting listed on the “RapCaviar” playlist with 8m. followers. Panels B and C show examples of older songs from a specific decade getting listed on the “All Out 80s” playlist with 6m. followers, and the “I Love My 90s Hip Hop” playlist with 4.5m. followers, respectively. Panels D-F show examples of songs getting listed on context-based playlists. Specifically, these panels shows example songs getting listed on the “Songs To Sing In The Shower” playlist (5.5m. followers), the “Broken Heart” playlist (2.9m. followers), and “Classic Road Trip Songs” playlist (2.9m. followers), respectively.

Figure 3 shows that in all 6 examples, the playlist listing coincides with a fairly sharp increase in the number of streams. When selecting these examples, we made sure that no other events that may provide an alternative explanation for the increase in streams occurred at the same time (e.g., album releases, TV advertising, radio airplay). This makes it more likely that the increase in streams is caused by the playlist listing.

While our data set contains more examples like these, where a single playlist listing within a specific time-window occurs, for the majority of songs and weeks, the attribution of the playlist effects is not as straightforward due to multiple listings. We will exploit examples like the ones in Figure 3 to create quasi-experiment-like settings with treated and control songs. To analyze this subset of the data, we will use a Difference-in-Differences estimator to establish an estimate that can serve as a reference for our main model.

Figure 3

Model free evidence



Notes. The black line shows the weekly streaming volume for each of the six example songs. The grey area indicates the listing duration of the respective songs on the playlist mentioned in the title.

Preliminary Quasi-Experimental Evidence

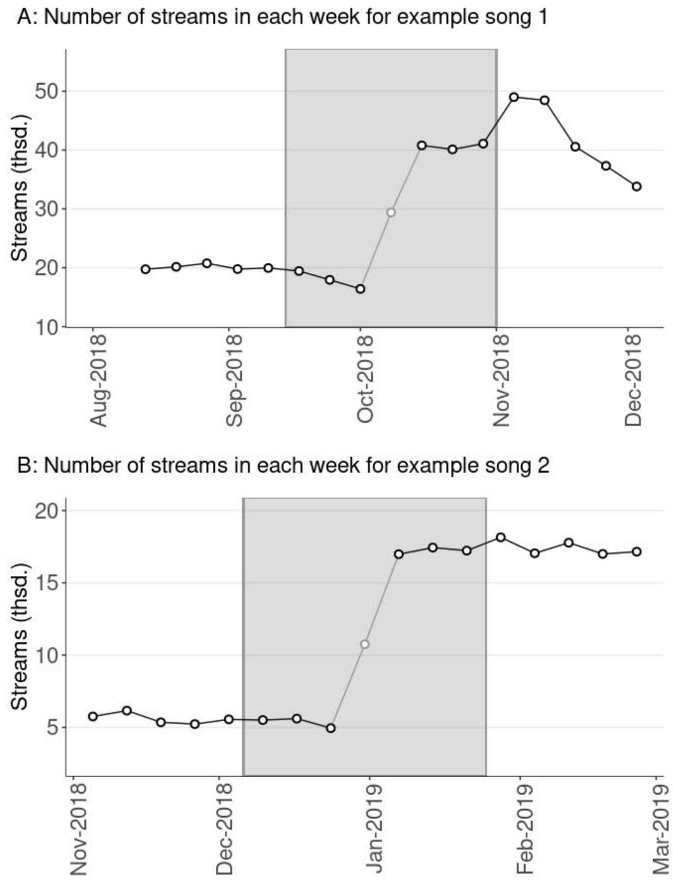
Ideally, we would be able to observe two very similar songs, with one song being placed on an additional playlist and a second "control" song not receiving this "treatment". In the absence of experimental data, we mimic this scenario by creating quasi-experiment-like settings. To this end, we identify all songs from our data that are placed on one or more additional playlists *once* in one week in a 7-week time window, i.e., songs qualify for our sample if they are first observed 3 weeks without changes in playlist listings, then a change in the number of playlist listings occurs in week 4, and after that the song is observed for 3 additional weeks without a change in playlist listings. These 3 "pre-treatment" weeks allow us to assess whether treated and control songs are similar prior to the playlist listing occurs, and the 3 post-treatment weeks allow us to assess the treatment's impact in the spirit of a quasi-experimental setting.

Figure 4 shows two example songs that meet the selection criteria for our sample. The grey areas indicate the 7-week windows we consider in our analysis, i.e., 3 weeks before and 3 weeks after the listing. The light grey sections of the lines indicate the week in which the song got listed on the playlist. We discard this week (week 4) because the change may occur at any day during this week so that we do not observe a full week including the playlist listing. The uplift due to the playlist listing is clearly visible in both examples.

In total, we identified 81,740 of these quasi-experiment-like settings, i.e., song-specific windows of 7 weeks with a change in playlists happening in week 4 and these 81,740 quasi-experiments cover 33,454 unique treatment songs, i.e., songs that get listed in one week, but do not have changes in their playlist listings 3 weeks prior and 3 weeks after the treatment. This means that 62% of all 54,020 songs in our sample are treated at least for one 7-week window and the remainder (38%) experience too frequent changes regarding the number of playlist listings to meet our selection criteria for treated songs. Note that a given song may have multiple 7-week time windows over the three 3-year observation period that qualify for an experiment according to the criteria described above. On average, a song from the group of treated songs qualifies for approx. 2.4 experiments and we assign unique experiment identifiers to each of these 7-week time

Figure 4

Time-windows of playlist listing for two example songs



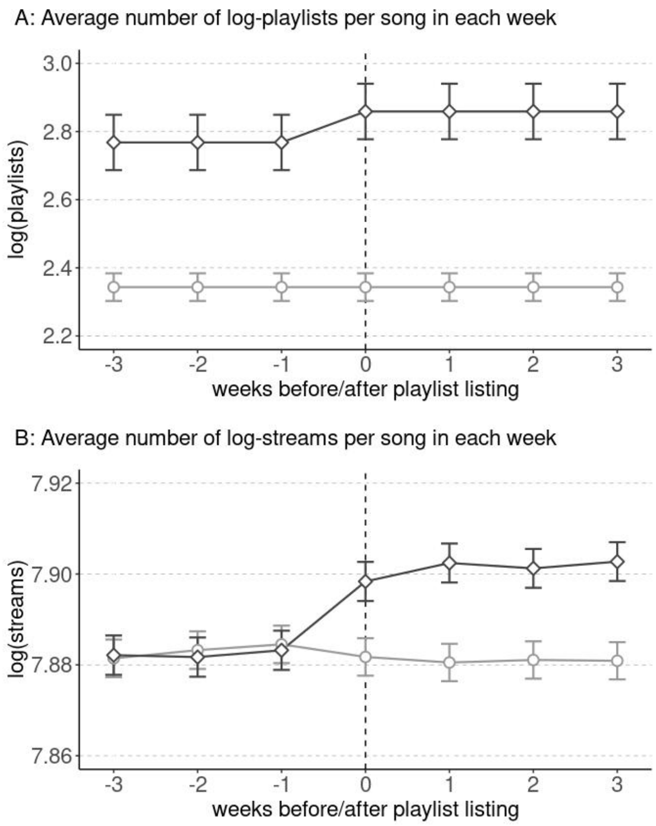
Notes. The grey area shows the considered time window for each example song in which the playlist listing occurred (i.e., three weeks before and after the playlist listing). The greyed out part of the line indicates the week when a song was included on the playlist and we exclude this week from the estimation since the day of the week when the song was added varies so that we don't observe a full week including the listing.

windows, which we denote *k* (*k*= 1,…, 81,740). Thus, if a song qualifies for multiple treatment windows, we assign a unique event ID for each window. Using fixed effects for each treatment window *k* in the analysis below allows us account for the fact that a given song's base-level of streams might change over time. To create a control group, we identify 98,070 control time-windows that are associated with 34,654 unique songs, i.e., 7-week time windows in which no changes in the playlist listings occur.

In Figure 5, Panel A, the black line shows the developments of the average number of playlist listings for the treated songs with the expected increase in playlist listings in week 4. The light grey line shows the control songs. Note that it is possible that a song gets listed on more than one additional playlist in week 4 of the 7-week period. The mean number of log-transformed playlist listings for the treated songs is 2.74 (sd = .95) before and 2.84 (sd = .86) after the additional listing(s) occur(s). This corresponds to non-transformed values of 23.2 (sd = 23.2) before and 24.4 (sd = 23.3) after the listing(s). Hence, on average songs receive 1.2 additional playlist listings. For the control group, Figure 5 shows that these values are just above an average of 2.16 (sd = .90) log-transformed playlist listings (i.e., non-transformed: 12.7, sd = 12.7) and, by construction, no change is observed over the observation window.

Figure 5

Control vs. treatment group in quasi-experiment-like settings



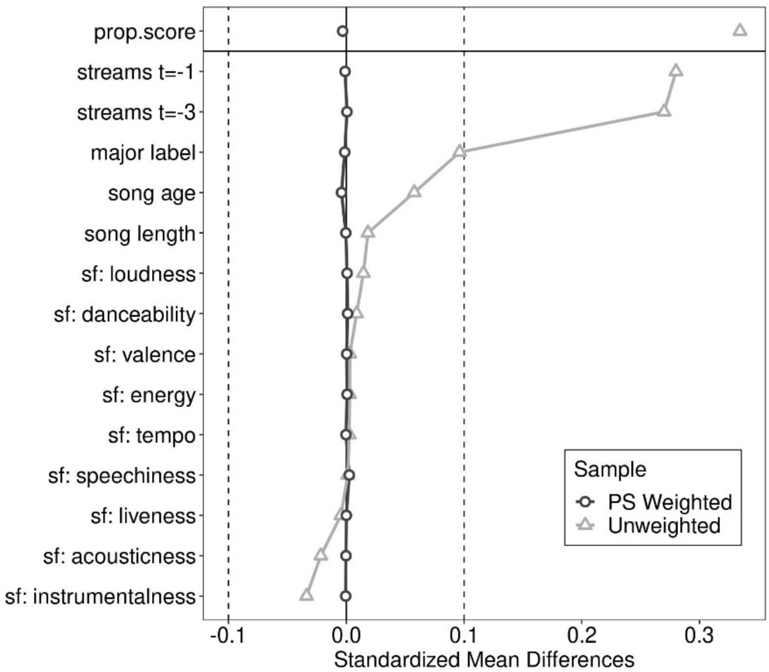
Notes. The dashed vertical line indicates the week when a song experiences an increase in playlist listings within the respective time window.

To make treatment and control songs as comparable as possible, we complement our analysis with a propensity score weighting approach (Inverse Probability of Treatment Weighting; Austin 2011). To this end, we estimate a logit model to obtain the probability of being treated:

(1) $$Pr(Treated_k = 1) = Pr(\beta_0 + Z_k\delta + \eta_k > 0) + \varepsilon_i,$$

where $Treated_k$ is a dummy variable indicating the treatment status of a given song within a respective 7-week time window, i.e., quasi-experiment k as described above. Z_k is a vector of matching variables that includes song characteristics that we obtained via the Spotify API (i.e., variables describing the musical style of a song such as valence, speechiness, and danceability), song length, song age, and the log-transformed number of streams in the first and last pre-treatment periods. Figure 6 shows that the matching adjusts very well for observable differences between treated and control songs. We report the results from the estimation of Eq. (1) and further details in Web Appendix B.

Figure 6
Covariate balance



Notes. The figure shows the standardized mean differences between the treated and control groups before matching (grey line) and after matching (black line). sf = song feature.

To quantify the main effect of playlist listings on demand, we estimate a weighted Difference-in-Differences (DiD) model of the following form, where log($Streams_{kst}$) is the log of the number of streams of song i in experiment k in experiment's week s, and $treatment_{kst}$ is a continuous treatment that is changed in period $s = 4$ of experiment k according to the change in the number of playlist listings:

$$log(Streams_{kst}) = \delta * treatment_{ks} + \pi_k + \tau_t + \phi_s + \varepsilon_{kst} \quad (2)$$

Because a song can potentially be part of multiple quasi-experiment-like events, we include treatment window fixed effects (π_k), which accounts for different baseline levels of streams of a song at different time periods during our 3-year observation period. In addition, we include week fixed effects (τ_t) to account for time-specific unobserved effects (e.g., seasonality), and a fixed effect for the week (ψ_s) that runs from 1-6 (remember we skip week 4 in the 7-week window). We implement the Inverse Probability of Treatment Weighting approach by estimating equation (2) with weighted least squares (Austin 2011, p. 409). We obtain the weights from propensity scores that we predict from equation (1), where $Pr(\text{Treated}_i = 1)_{ik}$ is the propensity score, i.e., the probability of being treated, given the observed covariates:

$$w_{ik} = treated_{ik} + \frac{\Pr(\text{Treated}_i = 1)_{ik}(1 - treated_{ik})}{(1 - \Pr(\text{Treated}_i = 1)_{ik})} \quad (3)$$

Figure 5 (Panel B) displays the average log-transformed number of streams for treated and untreated songs, and it shows that the songs have a very similar development prior to the treatment. This suggests that the assumption of parallel pre-treatments trends, which is critical for the identification of a causal effect, is likely to be fulfilled. To test whether the pretreatment trends are indeed statistically equivalent across the treated and non-treated groups, we carry out a "placebo" treatment test. That is, we restrict the sample to the 3 weeks before the treatment and define the treatment in the midpoint of these observations (i.e., week 2) (Datta, Knox, and Bronnenberg 2017). We then estimate a DiD model based on Eq. (2) on this data and fail to reject the null hypothesis of no treatment effect (δ = .001, p > .05, see Web Appendix C), indicating that the assumption of parallel pretreatment trends is met.

While the pre-treatment periods displayed in Figure 5 (Panel B) show a parallel development between treatment and control songs, the mean number of streams clearly deviate between these groups as of period 4, indicating the presence of a positive treatment effect. As can be seen from Table 4, estimating Eq. (2) on the full data using weighted least squares results in a positive and statistically significant effect. Specifically, we obtain a coefficient for a treatment effect of δ = .165, meaning that an increase in the number of playlists by 1% increases the number of streams by approximately .165%.

Table 4

Estimation results: model comparisons

	(1) *Log(streams)*	*(2)* *Log(streams/downloads)*	*(3)* *Log(streams/downloads)*
log(Number of playlists) (δ)	**.165***** (.016)	**.205***** (.013)	**.197***** (.006)
Fixed effects			
Song-Event	yes		
Event-week	yes		
Week	yes		
Song		yes	yes
Artist-week		yes	yes
R-squared	.989	.919	.863
F	517.88	8.81	17.24
p-value	.000	.000	.000
Number of Songs	33,454	33,454	54,020
Number of Events	81,740	-	-
Observations	1,078,860	1,078,860	7,275,833

Notes. Clustered standard errors in parenthesis; * p < .05, ** p < .01, *** p < .001.

This model specification has the advantage that it allows us to identify the causal effect of changes in playlist listings based on isolated playlist events, and it provides evidence that there is a substantial positive effect of being included on a playlist with an elasticity of .165. The disadvantage of this specification is that not all songs have time periods during which only one change in playlist listings occurs in a 7-week window. Especially the more successful songs and song-week combinations usually have changes every week with potentially multiple simultaneous listings and would therefore not be included in this quasi-experiment-like approach. Thus, the sample would be skewed towards less successful songs and song-week combinations, and we would miss out on analyzing the economically important songs and time periods. Further, the restricted sample also restricts us in building a rich set of theoretically and managerially interesting moderators and to gauge the long-term effects of playlist listings. We therefore use this analysis as a stepping stone to establish the presence of an effect and to create a point of reference for the effect size.

Main Model Specification

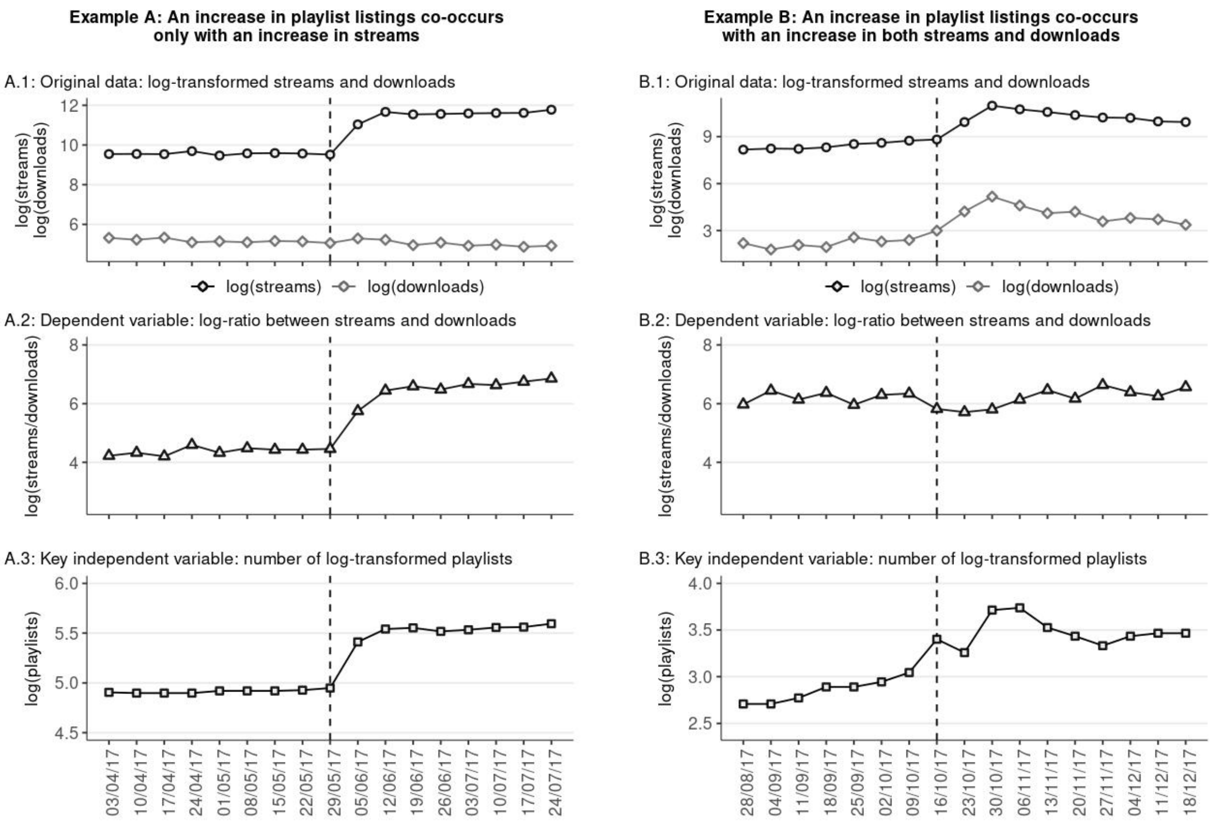
We now move on to the main econometric model that uses the full data set and relies on different identifying assumptions. As a first precaution against endogeneity we include song fixed effects. This safeguards against unobserved song differences causing it both to be included on playlists and having more streams. Further, because we observe multiple songs per artist, we include artist-time fixed effects to account for time varying shocks to artist popularity that may arise if, e.g., an artist releases a new album, goes on a concert tour, or appears in a TV show. However, there may be song-level demand shocks that vary over time that cannot be captured by these fixed effects. In particular, a song may become popular due to reasons not captured in our model, and this gain in popularity (streams) may lead to additional playlist listings.

We address this concern with our model specification that follows the spirit of Chevalier and Mayzlin (2006) by using as dependent variable the ratio between a song's streams and the song's paid downloads at time *t*. The rationale of this approach is that a song's playlist listings at time *t* should not causally affect the number of paid downloads this song receives. The reason is that playlists are limited to the streaming market, and it is highly unlikely that consumers *purchase* songs in a download store that they discovered on streaming playlists. At the same time, songs on streaming services are subject to the same underlying unobserved demand shocks.

Hence, if the correlation between playlist inclusion and streams is primarily driven by unobserved changes in popularity, the playlist listing should merely pick up unobserved demand shocks that are visible both in the streaming channel as well as in the download channel, and hence, the effect of playlists on the ratio (streams/downloads) should essentially be zero. However, if there is a true causal effect of playlist listings on streams, it should show up in a lift of the ratio (streams/downloads) when a song is added to a playlist. Based on these properties, this model specification allows us to estimate the effect of playlist on streams while controlling for unobserved demand shocks.

Figure 7

Illustration of identification approach using differencing



Notes. B.1 & B.2 show the developments of the log-transformed number of weekly streams and downloads for two example songs. A.2 & B.2 show the dependent variable in our main model, i.e., the log-ratio of the weekly number of streams and downloads (i.e., log($streams_{it}$) – log($downloads_{it}$)). A.3 and B.3 show the key independent variable in our model, i.e., the log-transformed number of playlists that the respective songs are listed on per week. The vertical dashed line indicates the weeks in which a change in the number of playlists for the respective songs occurred due to reasons that may be (partially) unobserved.

Figure 7 illustrates the intuition underlying our modeling approach using two example songs for selected time windows, within which an increase in playlist listings, our key independent variable, occurs for each song due to reasons that may be (partially) unobserved (Panels A.3 & B.3). Panel B.1 shows that for song B there is a simultaneous increase in the number of streams and downloads. For this song, our dependent variable (Panel B.2) differences out the variation that is due to an increase in the overall popularity of a song. The correlation between log(playlists) (Panel B.3) and the log(streams/download)-ratio (Panel B.2) is insignificant ($r = -.039$ $p = .879$), which means that the co-movement between playlist and streams observed for song B is not causal (and should not be used for causal inferences).

In contrast, for song A, there is no simultaneous increase in the number of downloads when it is added to a playlist, making it less likely that the effect of playlists on streams is confounded by unobserved demand shocks. In this case, the correlation between log(playlists) (Panel A.3) and the log(streams/download)-ratio (Panel A.2) is strong ($r = .984$, $p < .001$) and it does have a causal interpretation (and should be used for causal inferences). The identifying assumption that we have to make is that the unobserved time-varying song-specific shocks manifest themselves in both streaming services and download services and are not restricted to streaming services. We believe this is a reasonable assumption to make as (i) there are no readily available causes for these demand shocks in streaming services that are not already part of the analysis, and (ii) these shocks in the popularity of songs tend to be induced exogenously outside the system (e.g., a song becomes popular because of a viral video or commercial).

We estimate the following equation:

$$\text{(4)} \qquad \log\left(\frac{Streams_{ijt}}{Downloads_{ijt}}\right) = \alpha_1 PlaylistStock_{ijt} + \sum_{m=1}^{M} \beta_m \left(PlaylistStock_{ijt} * X_{m,\,ijt}\right) + \sum_{c=1}^{C} \gamma_c X_{c,\,ijt} + \delta_1\, AdStock_{ijt} + \delta_2 RadioStock_{ijt} + \delta_3 \log DownloadPrice_{ijt} + \eta_i + \mu_{jt} + \varepsilon_{ijt}$$

The dependent variable is the log of the ratio of streams divided by downloads for song *i* by artist *j* in week *t*. The focal independent variable is the log of the number of playlist listings for song *i* of artist *j* in week *t*. $X_{m,ijt}$ denotes a vector of moderating variables defined in Table 3. To allow for long-term effect of playlists, advertising and radio plays on a focal song's demand, we use a stock specification that directly yields their long-term elasticity (e.g., Dinner et al. 2014):

$$(5) \qquad \text{StockVariable}_{ijt} = \lambda_{Variable} * \text{StockVariable}_{ijt-1} + (1 - \lambda_{Variable}) * \left(ln\left(Variable_{ijt} + 1\right)\right)$$

We obtain values for lambda by maximizing fit using a grid search for values between 0 and .9 in increments of .1, leading to: $\lambda_{playlist} = .30$, $\lambda_{radio} = .60$, and $\lambda_{adspend} = .30$. In (4), η_i are the song fixed effects (for a total of 54,020 fixed effects) and μ_{jt} are the artist-time fixed effects (for a total of 1,910,291 fixed effects).

RESULTS

Establishing Validity

Before we proceed with the results for Eq. (4), we compare the results from the DiD analysis to a baseline model specification that only includes the log-transformed number of playlist listings as an explanatory variable and the song fixed effects and the artist-time fixed effects. This allows us to compare the results from these two model specifications with different identifying assumptions. Model (2) in Table 4 is restricted to the same subset of observations we used in the DiD specification but uses the log-ratio of streams and downloads as the dependent variable. The coefficient we obtain from this specification ($\delta = .205$, $p < .001$) is close the estimate from the DiD specification ($\delta = .165$, $p < .001$). The former result suggests that a 1% increase in playlist listings increases the number of streams by .205%.

Keep in mind that the specification in Model (1) is fairly strict, since this specification includes fixed effects for every 7-week time-window for each song. The coefficient in Model (2) is still close to the estimate in Model (1), despite this strict specification, and the confidence intervals of the estimates overlap (CI_{M1} = [.133; .196]; CI_{M2} = [.194; .215]). This strengthens our confidence that our main model specification allows for a causal interpretation of the coefficients. Table 4 also reports the findings for Model (3), which is the same as Model (2) but now uses all observations and again we obtain a coefficient that is comparable to the other models ($\delta = .197$, $p < .001$).[2]

Main Results

Table 5, Model (8) presents the results from our main model Eq. (4). We gradually build the full model by starting with a baseline specification that only contains the PlaylistStock variable in Model (4) and adding blocks of predictors for the main effects of the control, advertising and radio play, and moderating variables in Model (2), the interaction effects for the playlist characteristics in Model (6), and the interaction effects for the song and song-playlist variables in Model (7). To be able to interpret the direct effects to hold for an average observation, we grand means center variables before we calculate their product in the interaction terms.

The direct effect for playlist stock is $\alpha_1 = .349$, p < .001 and suggests that a 1% increase in the number of playlist listings leads to a .349% increase in the number of streams. We note that this magnitude is seemingly larger than for the models without moderating effects, but this is related to the grand mean centering of the moderating variables before they enter the interaction terms and their skewed distributions (Irwin and McLelland 2001). We will interpret the magnitude as a function of the different moderator variables based on Figure 8 below. We also

[2] In Web Appendix D, we show that a non-recommended alternative specification using log(streams) as the dependent variable ($\delta = .652$, $p < .001$) and a specification without artist-time fixed effects ($\delta = .747$, $p < .001$) produce results that suggest much stronger playlist effects, likely because they overestimate the playlist effect due to unobserved time-varying song-specific effects.

note that given the very large sample size of more than 7 million observations, it is not surprising that virtually all estimates are highly significant. As we will show, the effect sizes are considerable and economically substantive.

Moderation by playlist characteristics. The positive interaction effect between playlist stock with the *playlist followers* (β_1= .047, p < .001) indicates that the relative effect of an additional playlist listing is stronger the more followers a playlist has, which is as expected. Furthermore, commercial playlists appear to be slightly less effective in driving demand than playlist by independent curators, as evidenced by the negative interaction with the *commercial playlist share* variable (β_2 = –.003, p < .05). Although the magnitude of the effect is rather small, it suggests that trustworthiness, rather than expertise, affects consumer choices and brand consideration (Erdem and Swait, 2004), and independent entities are generally perceived as more trustworthy compared to commercial entities (Colicev et al. 2018).

With regard to the share of *context-based playlists*, we find a positive effect (β_3 = .020, p < .001), indicating that, all else equal, listings on playlists that are curated according to a specific context (e.g., running, commuting), rather than content-based criteria (e.g., genre, decade of release) will create a higher uplift in demand. We speculate that this finding may be due to the fact that the activities underlying the curation are reoccurring, and some even habitual, which ensures a regular interaction between the playlist followers and the listed content. In terms of *playlist length*, we find a negative interaction (β_4 = –.019, p < .001), suggesting that the more songs are co-listed on the same playlist, the less attention a song will receive, given the users' limited time-budget for music consumption.

Moreover, the *number of artists*, measuring how many different artists are co-listed on average with the focal artist, positively moderates the playlist effect (β_5 = .029, p < .001). This finding suggests that, as expected, songs from an artist benefit from being co-listed with a larger number of other artists on the same playlists through cross-promotion.

Moderation by song-playlist characteristics. We find that the variable measuring the *song's position on the playlist* negatively moderates the playlist effect ($\beta_6 = -.021$, $p < .001$), suggesting that the effect becomes less strong, the further down a song moves on a playlist's ranking of songs. Although the content on playlists can also be consumed in a random order, this finding suggests there is value to being listed on the top of the playlist. Furthermore, the effect of an additional playlist listing will be more positive, the more closely the type of music of the focal song matches the type of music of the other songs on the playlist, as measured by *song-playlist similarity* ($\beta_7 = .150$, $p < .001$). Although there may be potential benefits of being co-listed with dissimilar content, e.g., through exposure to new audiences (Oestreicher-Singer and Sundararajan 2012), our findings suggest that it is beneficial for content to be similar.

Moderation by song characteristics. We find that older songs benefit relatively more from playlist listings, as the interaction effect with *song age* shows ($\beta_8 = .110$, $p < .001$). This result is not surprising since new releases receive promotion via other channels (e.g., promotional videos, radio airplay), while older releases typically don't get this type of exposure and, thus, have a lower discoverability. This means that the relative effect of a playlist listing should have a stronger positive effect on product discovery for older releases, consistent with previous findings (Zhang 2018). With respect to the *artist fame* variable, we find that established artists benefit relatively less from promotion effects through playlists ($\beta_9 = -.045$, $p < .001$). This finding is consistent with our expectation that established artists capitalize on the audience they have built over their career, which should increase their discoverability vis-à-vis less-established artists who have not yet established a large fan base.

Table 5

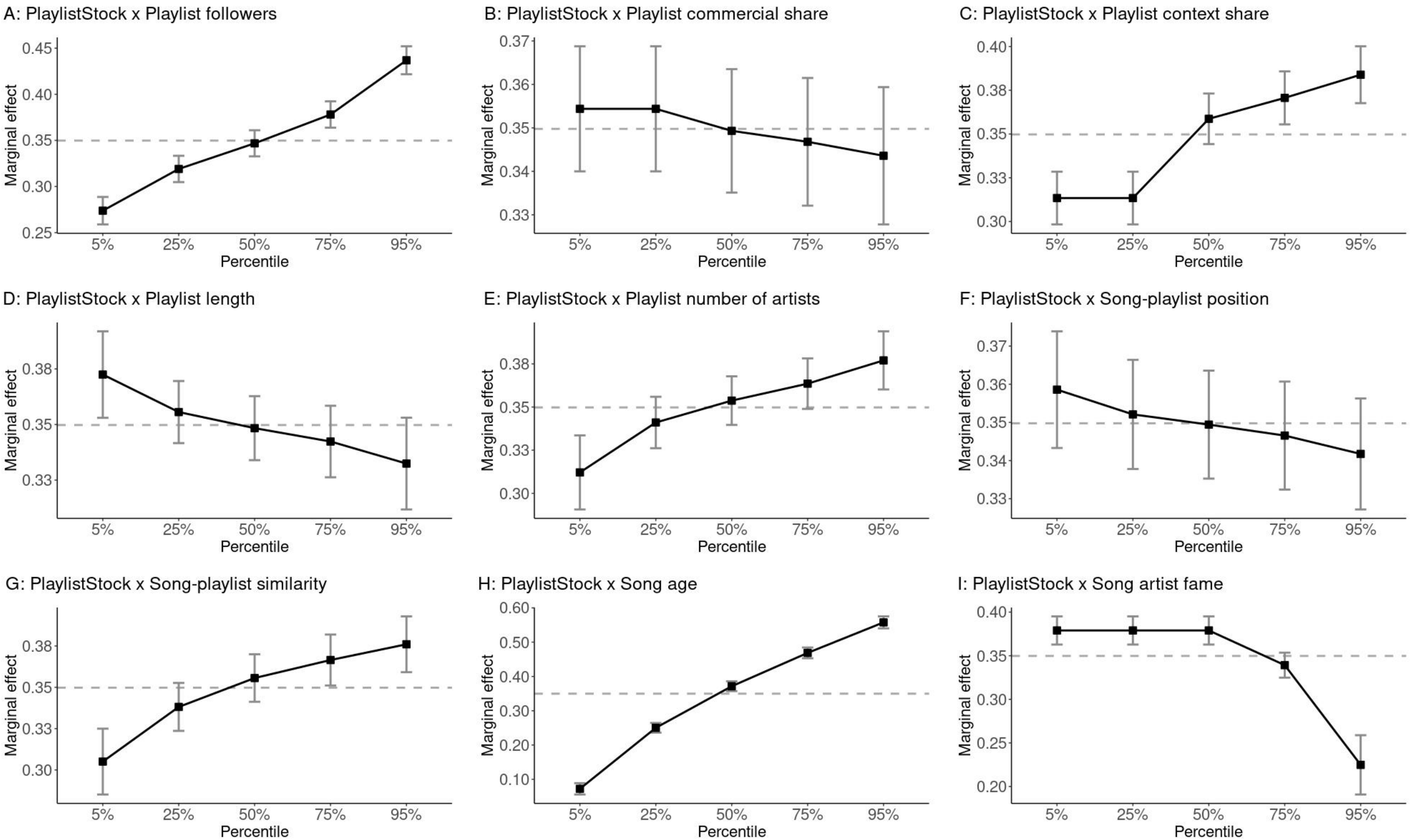
Main model results

Independent variables	Exp.	(4) Coeff. (SE)	(5) Coeff. (SE)	(6) Coeff. (SE)	(7) Coeff. (SE)	(8) Coeff. (SE)
PlaylistStock (α_1)	+	.202*** (.006)	.248*** (.006)	.307*** (.006)	.345*** (.006)	.349*** (.006)
PlaylistStock x Log(PlaylistFollowers) (β_1)	+			.040*** (.001)		.047*** (.001)
PlaylistStock x Log(PlaylistCommercialShare) (β_2)	+/–			–.013*** (.001)		–.003* (.001)
PlaylistStock x Log(PlaylistContextShare) (β_3)	+			.033*** (.001)		.020*** (.002)
PlaylistStock x Log(PlaylistLength) (β_4)	–			.102*** (.005)		–.019*** (.006)
PlaylistStock x Log(PlaylistNumberOfArtists) (β_5)	+			–.064*** (.013)		.029*** (.005)
PlaylistStock x Log(SongPlaylistPosition) (β_6)	–				–.026*** (.005)	–.021*** (.005)
PlaylistStock x Log(SongPlaylistSimilarity) (β_7)	+/–				.171*** (.022)	.150*** (.022)
PlaylistStock x Log(SongAge) (β_8)	+				.102*** (.002)	.110*** (.002)
PlaylistStock x Log(SongArtistFame) (β_9)	–				–.051*** (.005)	–.045*** (.006)
Log(PlaylistFollowers) (γ_1)			.066*** (.002)	.099*** (.002)	.089*** (.002)	.118*** (.002)
Log(PlaylistCommercialShare) (γ_2)			.008*** (.002)	–.000 (.002)	.005** (.002)	.008*** (.002)
Log(PlaylistContextShare) (γ_3)			.002 (.002)	.030*** (.002)	.000 (.002)	.018*** (.002)
Log(PlaylistLength) (γ_4)			.078*** (.007)	.188*** (.013)	–.045*** (.007)	–.013 (.013)
Log(PlaylistNumberOfArtists) (γ_5)			–.103*** (.005)	–.064*** (.013)	–.081*** (.005)	.013 (.012)
Log(SongPlaylistPosition) (γ_6)			.030*** (.007)	.037*** (.006)	–.034** (.011)	–.014 (.012)
Log(SongPlaylistSimilarity) (γ_7)			–.086 (.045)	–.007 (.046)	.169** (.053)	.150*** (.021)
Log(SongAge) (γ_8)			.108*** (.004)	.094*** (.004)	.179*** (.005)	.202*** (.005)
AdStock (δ_1)			–.041*** (.004)	–.039*** (.007)	–.036*** (.004)	–.037*** (.004)
RadioStock (δ_2)			–.066*** (.001)	–.064*** (.001)	–.054*** (.001)	–.058*** (.001)
Log(DownloadPrice) (δ_3)			.011 (.009)	.011 (.009)	.004 (.009)	.008 (.009)
R-squared		0.863	0.865	0.866	0.867	.868
F		17.24	17.61	17.73	17.84	17.93
p-value		0.000	0.000	0.000	0.000	.000

Notes: N = 7,275,833, number of songs = 54,020, max. number of weeks = 156; clustered standard errors in parenthesis; all interaction variables are centered around their grand mean; * $p < .05$, ** $p < .01$, *** $p < .001$; Stock variables: $\lambda_{playlist} = .30$, $\lambda_{radio} = .60$, and $\lambda_{adspend} = .30$; we check for multicollinearity in the full model specification and find that the VIF values are in an acceptable range (see Appendix 5).

Figure 8

Interaction effects



Notes. The marginal effects on the respective y-axes represent the elasticity of streams to the number of playlists.

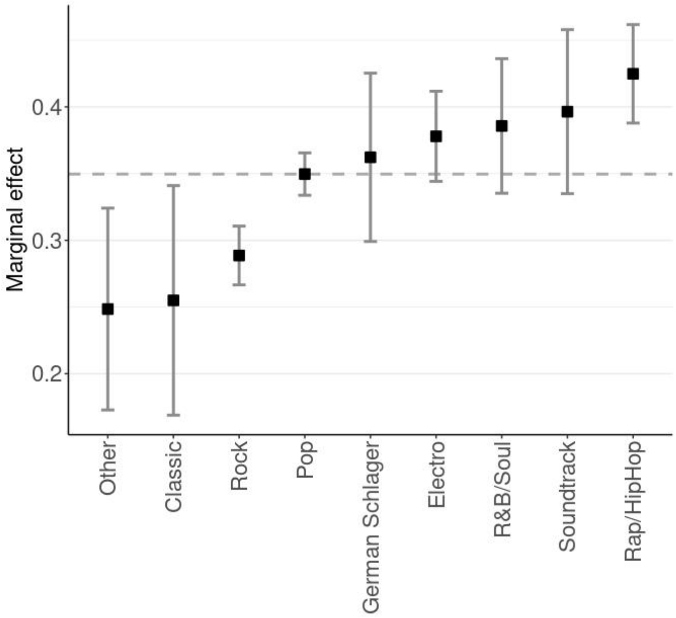
Advertising and radio airplay. We find a negative sign for *TV advertising* ($\delta_1 = -.037$, $p < .001$) and *radio airplay* ($\delta_2 = -.058$, $p < .001$), which may appear counterintuitive. However, since our dependent variable is the log-ratio of streams and downloads, this suggests that traditional marketing instruments are relatively less effective in access-based streaming channels compared to transaction-based download channels. In a separate analysis, we find that for streams (i.e., using log(streams) as the DV), the advertising elasticity is .008 ($p < .001$) whereas it is .045 ($p < .001$) for downloads (i.e., using log(downloads) as the DV). For streams, the radio airplay elasticity is .017 ($p < .001$) whereas it is .075 ($p < .001$) for downloads.

Figure 8 shows that the interaction effects imply substantial heterogeneity in the playlist effect across various dimensions, including song age, as visualized in Figure 8. The figure shows the marginal effect (elasticity) of the PlaylistStock variable on streams at the 5%, 25%, 50%, 75%, and 95% percentiles of the moderating variables. Inspecting the range of the elasticity across the different levels of the independent variables reveals substantial heterogeneity ranging from elasticities between close to 0 and values of .60.

Heterogeneity across Genres

We explore heterogeneity in the playlist effect across genres by adding an interaction term between the PlaylistStock variable and a categorical genre variable with 9 levels in Eq. (4), with Pop serving as a reference category. We visualize the genre-specific coefficients and their confidence intervals in Figure 9. This figure shows that the playlist effect is stronger for the Rap/Hip-Hop and R&B/Soul genres and less strong for the Classic and Rock genres. This may be due to the fact that Rap/Hip-Hop and R&B/Soul appeal rather to younger consumers who may be more open to integrating the new form of content curation through playlists into their listening behavior. We also find a slightly more positive effect for the Electro genre, which is similarly likely to attract a younger demographic. The relatively stronger positive effect for soundtrack songs can likely be explained by the nature of these compilations that feature multiple artists and promote exploration and discovery through the joint integration into the often emotional context of a movie. Overall, the genre effects appear face valid.

Figure 9
Playlist effect by genre



Notes. The marginal effect on the y-axes represents the elasticity of streams to the number of playlists with the Pop genre as the reference category.

Revenue Implications

To analyze the economic effect sizes of playlist listings, we use the estimated coefficients from our main model to predict the number of streams for different scenarios. In each scenario, we estimate the relative change in streams (i.e., the quasi-elasticity) resulting from one additional playlist listing on a playlist of size π for song *i* by artist *j* in week *t* ($\hat{\eta}_{ijt}$). We do so by predicting the number of streams based on a model where we add one playlist listing on a playlist of size π in week *t* for song *i* by artist *j* and dividing these predictions by the baseline predictions using the observed values in our data set. Because the size of a playlist is among the strongest drivers of playlist elasticity, we simulate outcomes for four different sizes of playlists, as represented by their mean number of followers: Top 5 playlists (π = 15,604,571 followers), Top 100 playlists (π

= 3,773,390 followers), Top 1,000 playlists (π = 927,730 followers), and Top 10,000 playlists (π = 41,811 followers). In the base case, a song is on average listed on 53 playlists, which becomes 53+1 after being added to one more playlist. We updated the average number of playlist followers for song *i* by artist *j* in week *t* according to the size of the playlists π. Thus, we use the following equation to compute the quasi-elasticity:

$$(6) \quad \hat{\eta}_{ijt} = \frac{\left(\widehat{Streams}_{ijt} \,\middle|\, \log(\exp(PlaylistStock_{ijt}) + 1), \log\left(\frac{SumPlaylistFollowers_{ijt}+\pi}{NumberOfPlaylists_{ijt}+1}\right)\right)}{\left(\widehat{Streams}_{ijt} \,\middle|\, PlaylistStock_{ijt}, \log\left(\frac{SumPlaylistFollowers_{ijt}}{NumberOfPlaylists_{ijt}}\right)\right)}$$

We compute the song-specific quasi-elasticity for song *i* by artist *j* by averaging across all observed weeks *t* for the respective song ($\bar{\eta}_{ij}$). We then multiply these song-specific quasi-elasticities by the average streams of the respective songs and by the average payout rate, i.e., the payments that streaming services pay to right holders per stream (i.e., \$0.006, Aguiar and Waldfogel 2021). Since most major streaming services and playlists are global platforms but our estimates of the change in streams only pertains to the focal market (Germany), we convert our estimates to the global level according to the German market share of the global recoded music market. Thus, we calculate the change in revenue as follows:

$$(7) \qquad \Delta\tilde{R}_{ij} = \frac{\bar{\eta}_{ij} * \overline{streams}_{ij} * payout}{MarketShareGermany}$$

Table 6 reports the results of the simulation exercise aggregated across all songs in our sample and the mean values for the respective sub-groups. We repeat this simulation for songs with different levels of success (i.e., Top 1,000 songs and Top 10,000 songs) to highlight the varying impacts for different types of songs. In addition, since song age is a strong moderator of playlist elasticity, we repeat the simulation for different subsets of the data that differ in terms of the release date. We follow industry standard and classify songs from their release up to 18 months after their release as new releases (so-called 'frontline' releases) and after that as catalog releases (Ingham, 2017). This 18-months cutoff is managerially relevant, since the release will also be handled by a different department of the music label. We report additional details regarding our prediction approach in Web Appendix F.

Inspecting the results for the full sample in Table 6 reveals that for an average song, a listing on a Top 10,000 playlist increases demand by 8.6% and the quasi-elasticities increase with the popularity of the playlist. For listing on a Top 5 playlist, the relative increase in demand is 47.4%. The associated monetary values show that the additional income through playlist listings increases with the audience size of the playlist. While the average song can expect to earn $130 additional revenue per week from a listing on a Top 10,000 playlist, a listing on a Top 5 playlist is worth an additional $789 per week. Keep in mind that the distribution of demand in our sample is highly skewed with the majority of songs being located in the tail of the distribution with relatively low demand. Hence, particularly the values for the Top 5 and Top 100 playlists need to be treated with caution, since we do not actually observe instances of relatively unpopular songs getting listed on these very popular playlists with millions of followers.

To consider a more realistic setting and avoid extrapolation beyond the observed playlist listings in our sample, we turn to the results of the Top 1,000 songs and Top 10,000 songs that are more likely to achieve a listing on the top playlists. It can be seen that while the average effect decreases for the group of Top 1,000 songs, the additional revenue per week increases substantially due to the higher baseline level of streams of these songs, e.g., to $7,196 per week for a Top 1,000 song getting listed on a Top 5 playlist. Note that the monetary effects described in the table refer to the additional revenue *per week*. Songs usually get listed on playlists for multiple weeks and the listing duration depends on the playlist type. In our sample, a song is listed on the Top 5 playlists for 17 weeks on average. Hence, the total effect for a playlist listing on a Top 5 playlist for a Top 1,000 song would amount to approximately $122,000, which is broadly in line with the estimates by Aguiar and Waldfogel 2021. The average listing duration on the less popular playlists is even longer (134 weeks) and this should be kept in mind when interpreting the dollar amounts.

We externally validate our estimates using public data from Spotify (*https://spotifycharts.com/*). During our observation period a song in the global Top 200 charts had approximately 8.4m. streams per week. Hence, using our quasi-elasticity estimates for the group of top songs (i.e., 18.1%) suggests that an additional listing on a Top 5 playlists would results in approximately 1.5m. additional streams for this group of songs. Using the payout rate of $.006 per stream suggests an additional weekly revenue of $9,000, which corresponds well to our estimate for the group of Top 1,000 songs (i.e., $7,196).

A pattern that emerges from Table 6, when comparing estimates across different success levels, is that for less popular a songs (i.e., songs with low streaming volume), we observe higher relative effects but lower absolute effects. This is expected, since the baseline level of streams in lower for the group of less successful songs. This pattern is also reflected when comparing new releases with catalog releases. It can be seen that while the relative effects are lower for newer (i.e., Frontline) releases the absolute monetary effects are higher compared to catalog songs. This finding also mirrors the interaction with song age that suggested higher elasticity for older releases that do not benefit from other promotional activities that usually accompany the releases of new songs.

Table 6

Revenue impacts

	All songs		*Top 10,000 songs*		*Top 1,000 songs*	
	quasi-elasticity	*avg. weekly Δ revenues*	*quasi-elasticity*	*avg. weekly Δ revenues*	*quasi-elasticity*	*avg. weekly Δ revenues*
All songs	# songs: 54,020		# songs: 10,000		# songs: 1,000	
Effect of 1 additional playlist listing on a…						
Top 5 playlist	47.4%	$789	34.4%	$1,568	18.1%	$7,196
Top 100 playlist	32.3%	$520	18.5%	$859	8.3%	$3,298
Top 1,000 playlist	20.2%	$315	8.9%	$423	3.2%	$1,295
Top 10,000 playlist	8.6%	$130	2.3%	$118	0.5%	$218
Frontline songs	# songs: 21,683		# songs: 5,701		# songs: 954	
Effect of 1 additional playlist listing on a…						
Top 5 playlist	32.1%	$1,424	25.8%	$1,597	15.2%	$8,091
Top 100 playlist	20.3%	$904	14.4%	$891	6.9%	$3,696
Top 1,000 playlist	11.9%	$529	7.3%	$457	2.7%	$1,442
Top 10,000 playlist	4.7%	$210	2.3%	$142	0.5%	$243
Catalog songs	# songs: 47,175		# songs: 6,342		# songs: 378	
Effect of 1 additional playlist listing on a…						
Top 5 playlist	50.7%	$497	39.8%	$1,542	27.0%	$4,939
Top 100 playlist	35.1%	$343	21.3%	$891	12.5%	$2,290
Top 1,000 playlist	22.2%	$217	10.1%	$457	5.1%	$924
Top 10,000 playlist	9.4%	$93	2.5%	$142	0.8%	$154

Notes: To compute revenues, we assume a payout rate of $0.006 per stream (Aguiar and Waldfogel 2021); the market share of Germany in $US trade value is ~7.6% according to aggregate market level statistics (i.e., $1.323b./$17.3b.; IFPI 2018); one song may be (partly) considered as a frontline song *and* (partly) as a catalog song if the song passes the 18 months cutoff for the classification during our observation period (i.e., the sum of the number of frontline songs and catalog songs is not equal to the overall number of songs in the respective samples); the average number of followers (π) for the different playlist types are 15,604,571 (Top 5), 3,773,390 (Top 100), 927,730 (Top 1,000), and 41,811 (Top 10,000), respectively; the average number of streams for the different song types are 17,882 (all songs), 516,011 (Top 1,000 songs), and 62,007 (Top 10,000 songs), respectively (details are provided in Web-Appendix F).

DISCUSSION

In this paper we have studied the effect of inclusion of a song on playlists on demand, and how this effect is moderated by various playlist and song characteristics. Using a music dataset unprecedented in coverage of not only large playlists but also small ones, and a wide range of more versus less successful songs, we estimate playlist elasticities and moderating effects. Based on analyses utilizing quasi-experimental data and a judicious econometric model (where the dependent variable corrects for otherwise unobserved demand shocks), we offer new empirical regularities for playlist elasticities based on a sample of more than 7 million song-week observations (covering approximately 54,000 songs from more than 13,000 artists observed across 156 weeks and more than 61,000 playlists). We now discuss implications for theory and practice as well as limitations and future research opportunities.

Theoretical Implications

The digitization of information goods such as music, books and video has completely changed the way consumers find, consume and pay for these goods. Whereas in the pre-digital era, consumers had to carefully consider their budgets given the relatively high prices of CDs, books and videos, today's access-based subscription services such as Spotify, Pandora and Netflix mean that consumers have access to a virtually unlimited amount of digital content at zero marginal cost. While this abundance is great news for consumers, it also creates a difficult choice task for consumers, which has also been referred to as the "tyranny of choice".

For the producers of information goods such as music artists, this new access-based business model comes with several challenges. One key challenge is that in the pre-digital era, producers were paid at the moment of purchase and did not have to worry about consumption, whereas nowadays, producer revenue is a function of how often digital content is accessed. For example, music artists only receive $.006 (less than 2/3rd of a dollar cent) per stream and hence it is in their interest to promote consumption. Since traditional forms of promotion such as TV advertising and radio airplay tend be to relatively ineffective, they have to find new ways to promote their content to audiences.

In this paper we posit that curation of digital information goods (e.g., through playlists) represents an important new marketing tool that addresses both the consumer's problem of abundant choice and a promotional vehicle for content producers to expose audiences to their digital content. We argue for conceptual analogies with brick-and-mortar retail distribution, which can be seen as a form of curation for physical products.

Building on these analogies, we derived and tested new expectations on what song and playlist characteristics make the inclusion of a song on a playlist more (versus less) effective for driving demand. We find that factors that enhance playlist exposure (the number of followers), the diversity of the exposure (more other artists), the relative visibility of the song on the playlist (near the top, shorter playlists) and fit (more similarity between song and playlist) enhance playlist elasticity. The way playlists are organized matters (stronger elasticity for organization by context) and by whom (commercially independent playlist are more effective). We also find that playlists offer a form of positive discrimination, in the sense that the playlists elasticities are stronger for less famous artists and older songs, where in both cases there is more upward potential for a playlist to lead to more streams. These moderating effects offer important new conceptual insights on the conditions in which playlist elasticities are weaker versus stronger. The finding that playlists benefit relatively unknown brands more is consistent with studies showing that demand will become more fragmented as new marketing tools, such as playlists, cater to niches that have previously been neglected due to a reduction in search costs (Brynjolfsson, Hu, and Simester, 2011). In a similar vein, Zhang (2018) shows that increased content sharing through the removal of digital right management restrictions on music download services benefits relatively unknown content disproportionally.

We study curation in the context of playlists in the music market, but we suggest that our findings are also relevant for actors in other industries. In other digital goods markets such as video content or books, playlists are not that common yet, probably because it takes longer to watch a TV show or movie or read a book that it takes to listen to a song and music is more prone to repeated consumption. However, other forms of curation have emerged that share similarities with playlists. In particular, personalized recommendations (based on a user's consumption history) are common and have received research attention (e.g., Adomavicius et al. 2018). Netflix has recently announced a new type of playlists for its users, called the Flixtape, which lets users group shows and movies according to mood. This development is in line with the context-based playlists we observe in our data. It would be worthwhile to study how playlists help shape consumer demand in other contexts with their own nuances.

Practical Implications

Digital content producers (e.g., music artists) and their agents (e.g., music labels) can use our findings for more informed decisions on playlist marketing. A first main takeaway is that (on an elasticity basis), playlists are much more effective in shaping demand than traditional forms of marketing and promotion such as TV advertising or radio airplays. This is not surprising since playlists are located “where the rubber hits the road” (i.e., within streaming contexts) whereas TV advertising and radio plays are located outside of these contexts. The inclusion of a song of playlist directly exposes subscribers to the song and gives them a free trial opportunity.

A key question is how artists and their management can identify or create the most influential playlists.

A first observation is that playlists are especially effective for less famous artists. However, many artists are not or they are completely independent. This is good news in particular for small artists who are not with a major music label. The results suggest they can directly work with independently managed playlists, which we have found are even more effective than commercial playlists.

To identify the most influential playlists, our further findings suggest prioritizing context-based playlists (that of course have to fit the artist’s music), shorter playlists with similar music yet more artists, and within these playlists aim for top positions on the list. Playlists with such characteristics will have elasticities easily in excess of 0.5 (given the size of the moderation effects in Figure 8), which means that 10% more playlist inclusions (e.g., going from 50 to 55 playlists) can lift demand and revenue by more than 5%.

Limitations and Opportunities for Future Research

While we document convergent results on playlist elasticities using different approaches to infer causal effects, one caveat is that it is impossible to rule out any remaining threats to causality with observational data. Ideally our findings would be tested in field experiments where songs are randomly added to playlists. To approximate the scope of our study, it would have to be a very wide-ranging experimental design, varying close to ten song- and playlist characteristics. Based on our discussion with several representatives of major music labels, it is extremely unlikely that such field experiments would ever be conducted. That is why believe that our analysis (built up gradually from preliminary quasi-experimental evidence to a rich econometric models) is possibly as close to causality as one can get for our research questions.

It would also be interesting to study how individual users differ in their use of playlists and the effect on the volume and variety of their (music) consumption. Follow-up research can also look at the role of playlist curators. From a marketing perspective, playlist curators can be viewed as influencers, i.e., persons with a large and engaged follower base (Haenlein et al. 2020).

We hope that this research will stimulate new thought and research into the role of curation. With the rise of digital information goods and access-based consumption there are ample new marketing research questions around curation that need to be addressed.

References

Adomavicius, Gediminas, Jesse C. Bockstedt, Shawn P. Curley, and Jingjing Zhang (2018), "Effects of Online Recommendations on Consumers' Willingness to Pay," Information Systems Research, 29 (1), 84–102. Ailawadi, Kusum L. and Paul W. Farris (2020), "*Getting Multi-Channel Distribution Right*," Wiley: New Jersey.

Aguiar, Luis and Joel Waldfogel, (2021), "Platforms, Power, and Promotion: Evidence from Spotify Playlists," *Journal of Industrial Economics*, to appear.

Anderson, Chris (2004), "The Long Tail," *Wired Magazine*, 12(10), 170–77.

Ataman, M. Berk, Carl F. Mela, and Harald J. Van Heerde (2008), "Building Brands," Marketing Science, 27 (6), 949-1136.

Ataman, M. Berk, Harald J. Van Heerde, and Carl F. Mela (2010), "The Long-Term Effect of Marketing Strategy on Brand Sales," *Journal of Marketing Research*, 47 (5), 866-82.

Austin, Peter C. (2011), "An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies," *Multivariate Behavioral Research*, 46 (3), 399–424.

BBC (2021), "Fifth of all songs streamed in UK in 2020 were rap and hip hop," *bbc.com*, https://www.bbc.com/news/newsbeat-56749586 (accessed: 05-09-2021).

Brynjolfsson, Erik, Yu (Jeffrey) Hu, and Duncan Simester (2011), "Goodbye Pareto Principle, Hello Long Tail: The Effect of Search Costs on the Concentration of Product Sales," *Management Science*, 57 (8), 1373–86.

Brynjolfsson, Erik, Yu (Jeffrey) Hu, and Michael D. Smith (2010), "Long Tails vs. Superstars: The Effect of Information Technology on Product Variety and Sales Concentration Patterns," *Information Systems Research*, 21 (4), 736-47.

Campo, Katia, Els Gijsbrechts, Tom Goossens, and Ann Verhetsel (2000), "The Impact of Location Factors on the Attractiveness and Optimal Space Shares of Product Categories," *International Journal of Research in Marketing*, 17 (4), 255–79.

Chernev, Alexander, Ulf Böckenholt, and Joseph Goodman (2015), "Choice Overload: A Conceptual Review and Meta-Analysis," *Journal of Consumer Psychology*, 25 (2), 333–58.

Chevalier, Judith A. and Dina Mayzlin (2006), "The Effect of Word of Mouth on Sales: Online Book Reviews," *Journal of Marketing Research*, 43 (3), 345-54.

Colicev, Anatoli., Ashwin Malshe, Koen Pauwels, and Peter O'Connor (2018), "Improving Consumer Mindset Metrics and Shareholder Value through Social Media: The Different Roles of Owned and Earned Media," *Journal of Marketing*, 82 (1), 37-56.

Datta, Hannes, George Knox, and Bart J. Bronnenberg (2017), "Changing Their Tune: How Consumers' Adoption of Online Streaming Affects Music Consumption and Discovery," *Marketing Science*, 37 (1), 1-175.

Dinner, Isaac M., Harald J. van Heerde, and Scott A. Neslin (2013), "Driving Online and Offline Sales: The Cross-Channel Effects of Traditional, Online Display, and Paid Search Advertising," *Journal of Marketing Research*, 51 (5), 527-45.

Elberse, Anita (2010), "Bye-Bye Bundles: The Unbundling of Music in Digital Channels," *Journal of Marketing*, 74 (3), 107-23.

Eliashberg, Jehoshua and Steven M. Shugan (1997), “Film Critics: Influencers or Predictors?,” *Journal of Marketing*, 61 (2), 68-78.

Erdem, Tülin and Joffre Swait (2004), “Brand Credibility, Brand Consideration, and Choice,” *Journal of Consumer Research*, 31 (1), 191-98.

Goh, Khim-Yong, Cheng-Suang Heng, and Lin Zhijie (2013), “Social Media Brand Community and Consumer Behavior: Quantifying the Relative Impact of User- and Marketer-Generated Content,” *Information Systems Research*, 24 (1), 88-107.

Haenlein, Michael, Ertan Anadol, Tyler Farnsworth, Harry Hugo, Jess Hunichen, and Diana Welte (2020), “Navigating the New Era of Influencer Marketing: How to be Successful on Instagram, TikTok, & Co,” *California Management Review*, 63 (1), 5–25.

Hendricks, Ken and Alan Sorensen (2009), “Information and the Skewness of Music Sales,” *Journal of Political Economy*, 117(2), 324–69.

IFPI (2018), “IFPI Global Music Report 2018”, London.

Ingham, Tim (2017), “If Daniel Ek Wants to Frighten the Major Labels, He Should Step Back in Time,” *musicbusinessworldwide.com,* https://www.musicbusinessworldwide.com/if-daniel-ek-wants-to-frighten-the-majors-he-should-head-back-in-time/ (accessed: 05-09-2021).

Irwin, Julie, R. and Gary H. McClelland (2001), "Misleading Heuristics and Moderated Multiple Regression Models," *Journal of Marketing Research*, 38 (1), 100-09.

Lamberton, Cait Poynor, and Kristin Diehl (2013), “Retail Choice Architecture: The Effects of Benefit- and Attribute-Based Assortment Organization on Consumer Perceptions and Choice,” *Journal of Consumer Research*, 40 (3), 393–411.

Liu, Hongju, Qiang Liu, and Pradeep K. Chintagunta (2017), “Promotion Spillovers: Drug Detailing in Combination Therapy,” *Marketing Science*, 36 (3), 382–401.

Mulligan, Mark (2015). Why Niche Is The Next Streaming Frontier. *Music Industry Blog*. https://musicindustryblog.wordpress.com/2015/09/04/why-niche-is-the-next-streaming-frontier/ (accessed: 05-09-2021)

Mulligan, Mark (2019). Global Streaming Music Subscription Market, H1 2019. *midiaresearch.com*. https://www.midiaresearch.com/blog/music-subscriber-market-shares-h1-2019 (accessed: 05-09-2021)

Natan, Olivia R. (2021), *Choice Frictions in Large Assortments*. Working Paper.

Oestreicher-Singer, Gal and Arun Sundararajan (2012), “The Visible Hand? Demand Effects of Recommendation Networks in Electronic Markets,” *Management Science*, 58 (11), 1963–81.

Rooderkerk, Robert P. and Donald R. Lehmann (2020), “Incorporating Consumer Product Categorizations into Shelf Layout Design,” *Journal of Marketing Research*, 58 (1), 50-73.

Rosario, Ana Babić, Francesca Sotgiu, Kristine De Valck, and Tammo H.A. Bijmolt (2016), “The Effect of Electronic Word of Mouth on Sales: A Meta-Analytic Review of Platform, Product, and Metric Factors,” *Journal of Marketing Research*, 53 (3), 297-318.

Sahni, Navdeep S. (2016), “Advertising Spillovers: Evidence from Online Field Experiments and Implications for Returns on Advertising,” *Journal of Marketing Research*, 53 (4), 459–78.

Scheibhenne, Benjamin, Rainer Greifeneder, and Peter M. Todd (2010), "Can There Ever Be Too Many Options? A Meta-Analytic Review of Choice Overload," *Journal of Consumer Research*, 37 (3), 409–25.

Salganik, Metthew J., Dodds, Peter Sherdian, and Watts, Duncan J. (2006), "Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market," *Science*, 311 (5762), 854–856.

Schwarzt, Barry (2004a), “The Tyranny of Choice,” *scientificamerican.com*, https://www.scientificamerican.com/article/the-tyranny-of-choice/ (accessed: 05-09-2021).

Schwarzt, Barry (2004b), “The Paradox of Choice: Why More Is Less,” Harper Perennial.

Sethuraman, Raj, Gerard J. Tellis, and Richard A. Briesch (2011), “How Well Does Advertising Work? Generalizations from Meta-Analysis of Brand Advertising Elasticities,” *Journal of Marketing Research*, 48 (3), 457-71.

Shankleman, Jess (2021), “Avoiding the Tyranny of Choice,” https://www.bloomberg.com/news/articles/2021-02-26/avoiding-the-tyranny-of-choice-on-tinder-on-netflix-or-at-work (accessed: 05-09-2021).

You, Ya, Gautham G. Vadakkepatt, and Amit M. Joshi (2015), “A Meta-Analysis of Electronic Word-of-Mouth Elasticity,” *Journal of Marketing*, 79 (2), 19–39.

Zhang, Laurina (2018), “Intellectual Property Strategy and the Long Tail: Evidence from the Recorded Music Industry,” *Management Science*, 64 (1), 24-42.