European Commission

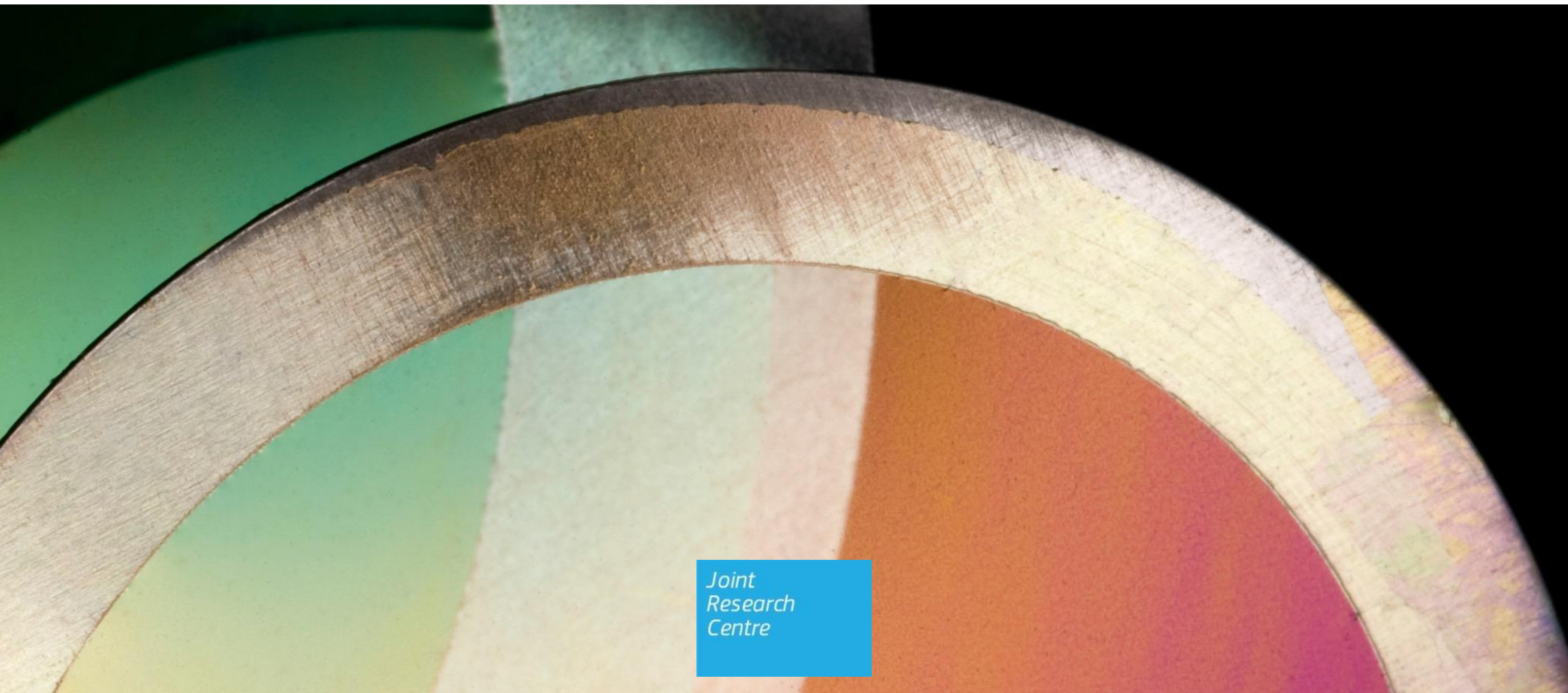
JRC TECHNICAL REPORTS

JRC Digital Economy Working Paper 2018-07

Playlisting Favorites: Is Spotify Gender-Biased?

Luis Aguiar
Joel Waldfogel (University of Minnesota and NBER)
Sarah Waldfogel (University of Wisconsin)

November 2018



JRC Science Hub
https://ec.europa.eu/jrc

JRC113503

Seville, Spain: European Commission, 2018

How to cite: Luis Aguiar, Joel Waldfogel, Sarah Waldfogel *Playlisting Favorites: Is Spotify Gender-Biased?* JRC Digital Economy Working Paper 2018-07; JRC Technical Reports, JRC113503.

Abstract

The growth of online platforms has raised questions about their power and the possibility that it could be exercised with bias, including by gender. Women account for about a fifth of the most successful artists at Spotify, prompting some concerns about bias. We explore the roles of female participation, along with promotion decisions at the platform - in particular playlist inclusion - in explaining the female share of successful songs and artists at Spotify in 2017. We employ two broad tests for gender bias. First, we ask whether songs by female artists are differently likely to appear on global playlists, conditional on the past success of the artists, song characteristics such as genre, and gender. Second, we test for bias in New Music Friday playlist ranking decision based on outcomes, asking whether songs by female artists stream more, conditional on their New Music Friday rankings. We find some evidence consistent with bias (in favor of women at Today's Top Hits as well as in the New Music rankings, and against women at some global playlists). These biases, however, do little to explain the low female share of streaming on Spotify, which we instead attribute to the relatively low share of female songs entering the platform.

Contents

Acknowledgments

We thank participants in the European Commission's Joint Research Center Digital Economy Seminar held in Seville on September 21, 2018.

1 Introduction

The rise of online platforms has raised questions about their possible power to influence market and other social outcomes, as well as whether their operators exercise their power in a biased fashion.[1] These issues arise in the music market as well as many others. Prior work documents the power of Spotify to influence song success through its editorial playlist decisions (Aguiar and Waldfogel, 2018). For example, inclusion on Spotify's Today's Top Hits playlist adds about 20 million streams to a song's total. In the wake of widespread revelations of sexual harassment in the motion picture industry and the #MeToo movement, the recorded music industry has also come under scrutiny for its treatment of women. Recent popular and academic studies find that while women make up roughly half of audiences, female artists account for less than a quarter of successful performers and an even smaller share of successful songwriters.[2] Some observers have raised concerns about bias, exhorting platforms like Spotify to do more to promote female artists, while others have encouraged more female participation in the industry.

The relative low female share among the most successful songs has, broadly speaking, two possible causes. First, like many occupations with unrepresentative participation by gender, musical performance may attract more male than female participation. In that case, a low female share of song sales or streams might stem from factors prior to promotion decisions by platforms such as Spotify.[3] Second, women may face bias in the support they receive from intermediaries, including record labels, radio stations, or streaming services, whose decisions on investment and promotion can have major effects on what succeeds. Determining which of these explanations are relevant is important, as they suggest different remedies.

Digitization has reduced barriers to entry in the cultural industries. New technologies have sharply reduced the cost of bringing works to market in music, movies, television, and books, bringing about explosions of new products. The reduction in barriers to entry means that all kinds of artists can bring a product to market, and artists added nearly a million songs to Spotify in 2017. But access may not be enough. First, artists of varied demographic backgrounds need to create music and make it available. Second, the platforms need to

[1] See Edelman (2011); Zhu and Liu (2016); Lambrecht and Tucker (2018).

[2] See Smith et al. (2018), Pelly (2018), and Flanagan (2018). Similar concerns have been raised in the film industry. See, for example, Ellis-Petersen (2014).

[3] The questions posed here about Spotify are similarly relevant to other streaming services such as YouTube, Deezer, and Apple Music. Only Spotify makes streaming data available.

provide commensurate levels of promotion to songs by female and male artists, given the strong effect of platform promotion decisions on artist and song success.

The goal of this study is to ascertain the determinants of female artists' success, and to assess the relative roles of female participation, and possible platform gender bias, on the success of female artists on the Spotify platform. We begin with a few purely descriptive questions. First, what are the gender shares of songs entering the platform and top songs streaming on the platform? Second, what share of songs on the Spotify-owned major playlists are by female artists? While these baseline facts provide a strong suggestion that women have entered the recorded music industry at lower rates than men, they do not rule out other explanations for the low female streaming shares among successful songs. We then attempt to test for bias in playlist inclusion - or playlist ranking decisions - using two broad approaches. Using what we term the "conditioning on observables" approach, we ask how playlist inclusion, or playlist rankings, depend on song and artist characteristics. After accounting for plausible determinants of whether songs are included in playlists - for example, past artist streams, artist genre, and song characteristics - we ask whether songs by female artists are more or less likely to be included, or to receive more favorable placement, on playlists. Our second test for bias, which we term the "outcome-based test," looks for an ex post symptom of bias on the New Music Friday lists. If songs by female artists receive less favorable New Music Friday playlist rankings than the songs should have, then those songs' streaming should exceed the streams for songs by non-female artist, holding playlist rank constant.

We create three broad datasets for this study. First, we assemble data on the songs entering Spotify between mid-2015 and early 2018. For the solo artists, we mechanically match the first names of first-listed artists with data on the names' gender frequencies as a way to measure the gender composition of songs entering Spotify. Second, we create a dataset of songs whose first or second listed artist is among the top 200 artists, by annual streams, in at least one of the 26 sample countries. Combining these top 200 artists and the other artists on songs including these top artists gives a total of 2,284 distinct artists, whose genders we ascertain by hand. These songs account for 95.4 percent of the streams reported by Spotify (the daily top 200 tracks by country). We refer to this as the "top artist" dataset. This allows us to measure the gender composition of songs and streaming among the top artists. We then combine this top artist dataset with the list of songs making seven major global playlists between May 1 and December 31, 2017 so that we can study the determinants, and possible gender biases in making these playlists. This dataset includes 6,650 songs. Third, we construct a list of songs ranked 20 or better on the New Music playlists in 26 countries during 2017. We obtain measures of artist genders, and we combine these data with song and artist characteristics, as well as eventual streaming success of these songs appearing on the New Music lists.

Ascertaining the genders of a song is a bit involved, but given its centrality in this study we introduce our approach here. While it is a relatively simple matter to determine the gender for a song by a solo artist, many songs are either by multi-person bands or are created by more than one listed entity (where the entity might be either a solo artist or a band). We code individual artists as male or female if we can ascertain gender (or as unknown if we cannot). We code bands as male or female if all prominent members (e.g. those included in official band photographs) are of the same gender, as mixed if the band includes both men and women, and as unknown if we cannot ascertain genders. We then code the gender of the song according to the gender information for the first two listed artist entities, using a sequence of possible approaches, from more to less restrictive. Our most restrictive measure treats a song as female only if its sole artist, or both entities (artists or groups), are entirely female. We refer to this as the "all female" measure. Our second measure treats the song as female if its sole artist is female, or in the case of two artist, both entities are either female or mixed ("female or mixed"). Our third measure deems the song female if the first entity is female regardless of the possible second ("first artist female"). Our most liberal measure classifies the song as female if either the first or second entity is female ("either artist female").

We find that among songs whose first listed artist is an individual (rather than a group), those who are female account for about a fifth of the songs entering Spotify. Songs by female artists account for between 13.7 and 22.6 percent of songs in the top artist sample, depending on which gender measure we employ. Songs with female involvement account for varying shares on the major playlists. Using the first artist female measure, the share varies between 4.0 percent on Rock This and 29.7 percent on Today's Top Hits. Using both conventional regression approaches as well as machine-learning based approaches to variable selection (as described in Belloni et al., 2014), we find that songs by female artists are substantially more likely to appear on Today's Top Hits. Songs by female artists are less likely to appear on

RapCaviar.[4] Controlling for observables in similar ways, we find that songs by female artists receive more favorable New Music Friday playlist ranks. Results on whether these more favorable ranks reflect gender bias are mixed: using the more restrictive female measures, we find evidence of favoritism in that female songs stream less. Using the less restrictive gender measures, we find no evidence of favoritism. We investigate the role of the biases we document for global playlists and conclude that the female share of successful songs at Spotify mainly arises from the relatively low female share of songs on the platform rather than anti-female bias in playlist decisions.

This study proceeds in five sections after the introduction. Section 2 presents a framework highlighting the possible explanations for the female share of successful recording artists, along with a discussion of existing evidence. Section 3 describes the data used in the study, including the process employed for determining the female role in particular songs. Section 4 presents descriptive findings on the female shares of songs entering Spotify and songs streaming on the service. Section 5 then turns to tests for bias in Spotify playlists inclusion decisions. We first present "conditioning on observables" tests for gender bias in both global curated playlists, as well as the New Music Friday lists. Section 5 also presents results of outcome-based tests for gender bias in the New Music Friday playlists ranks. Section 6 concludes.

2 Background

2.1 Framework

Explaining the female share of successful music on Spotify requires some explanation of the process by which music comes to market and sometimes succeeds. Succeeding in music requires an artist - and their product - to proceed through various junctures. Even if digitization has eliminated barriers to producing and distributing music, three potentially significant barriers remain. First, potential artists need to decide to create music for distribution in the first place. The burden of past discrimination may deter artists of various types from creating or recording music in the first place. Musician may be among the many

[4] We also find that female songs are less likely to appear on Rock This, but our top artist sample contains only 25 songs that appear on Rock This, giving us pause about drawing strong conclusions.

occupations that are disproportionately pursued by men (Hegewisch and Hartmann, 2014).

We note, however, that "explaining" a low female share of commercially successful music by a low female participation in the music industry would not make that low rate any more socially desirable. Rather, it would point toward different remedies.

The Bureau of Labor Statistics (BLS) maintains data on participation in occupations by gender.[5] Table 1 reproduces US data for 2018 on the demographic distribution of arts and media occupations. The female shares of these occupations vary widely, from 12.2 percent of the persons employed as "television, video, and motion picture camera operators and editors," to 60.9 percent for "writers and authors." "Musicians, singers, and related workers" are 37.7 percent female. It is difficult to know exactly how the BLS category of musicians and singers maps into would-be popular music creators, but these data do indicate that female participation in this occupation lags behind male participation.

Second, given the enormous amount of new recorded music available, intermediaries' decisions about what to promote can have an important influence on which music - and which artists - have an opportunity to succeed. There are many such gatekeeping intermediaries, including record labels that control decisions about whom to sign and promotional budgets for the products they release, as well as radio program directors and streaming platforms whose airplay and playlist decisions effectively determine which songs consumers hear. The intermediaries that we study are the playlist editors and algorithms at Spotify, whose decisions determine which music - and whether male or female-created music - gets promoted to listeners. Aguiar and Waldfogel (2018) document the important role of major Spotify playlists in determining song and artist success. Here we explore whether decisions to promote particular songs and artists may be made differently for artists of different genders.

2.2 Existing evidence

This paper contributes to two existing literatures. First, there is a growing body of work on possible platform bias. For example, Edelman (2011) explores whether Google biases its search results in favor of its own properties, and Zhu and Liu (2016) study whether Amazon

[5]Similar occupational segregation exists in Europe, although data on musicians in particular are not available. See https://ec.europa.eu/info/sites/info/files/150119_segregation_report_web_en.pdf, page 50 or http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=cult_emp_sex&lang=en.

enters the markets for products established by its marketplace vendors. Lambrecht and Tucker (2018) investigate the role of algorithms in possible gender bias of ad targeting.

Second, this paper contributes to ongoing discussions of gender bias in the music industry. For example, Smith et al. (2018) analyze data on "600 popular songs on the Billboard Hot 100 year end charts from 2012 to 2017," and conclude that "women comprised just 22.4% of artists and 12.3% of songwriters." The study finds "lack of women in these roles is surprising, given that women are a powerful market for music consumption: women comprised 53% of digital music buyers in 2014."

Pelly (2018) finds, based on a month of monitoring some popular Spotify playlists (including Today's Top Hits, New Music Friday, Rock This, RapCaviar, Hot Country, and ¡Viva Latino!) that many were "staggeringly male-dominated." More specifically, Pelly found that on Today's Top Hits, "64.5 percent of the tracks were by men as the lead artist, with 20 percent by women and 15.5 percent relying on collaborations between men and women artists." Existing work calls attention to low female shares of artists among successful and among heavily promoted songs. It would be useful to try to distinguish among possible explanations of the level of female success and promotion. Is it attributable to female participation in music or promotion by platforms? We now turn to data analysis geared to provide some answer to these questions.

3 Data

Analysis in this paper draws on four distinct underlying datasets. First, we have the full list of songs added to Spotify between late 2015 and early 2018, from Everynoise.com.[6] Some artist names are bands, others are individuals. By matching first names of the first listed artist with name gender databases, we can infer genders.[7] We are able to match genders for roughly 30 percent of first-listed artists.

Our second dataset is a group of songs at risk of making the major global Spotify playlists in the study. To study the determinants of being put on the major playlists, we begin with

[6] Everynoise.com is a site maintained by Glenn McDonald, a data scientist at Spotify. The site lists all songs entering Spotify and provides other information on streaming at Spotify.

[7] We use the U.S. Social Security names database as well as the website https://gender-api.com, which allows to determine gender based on a first name and covers non-US names.

the set of songs observed streaming in the daily top 200 on any day during 2017 in any of our 26 countries.[8] This is a large set of songs with a large set of artists, too many artists for us to determine all of their genders. To undertake the dataset creation process in a feasible way, we determine a set of artists, those appearing among the top 200 first-listed artists according to their country-specific streams during 2017. If there were no overlap in these "top artists" across countries, then this list would include 5,200 artists. There is, however, a good deal of overlap, so the total number of artists in this group is 1,695.

Our major playlist data cover the period beginning May 1, 2017. To make the sample for the analysis of whether songs make the playlists compatible with the playlist data we have, we restrict attention to the streams for songs first appearing in the streaming data on or after May 1, 2017. Because songs can have multiple artists, determining the gender of a song - and determining which songs to include is a bit complicated. While some songs include many artists, the vast majority of songs include two or fewer artists. Our analysis sample includes all of the songs on which either of the first two artists was among the original "top artist" sample. This required us to find the genders, as above, for an additional 963 artists.

We determine the genders of each of the 2,758 first or second-listed artists by hand, generally using one of the following sources: the Spotify artist page, Wikipedia, Allmusic, and press mentions. We relied on the pronouns used for describing artists as well as photographs. Our measures of gender are therefore best thought of as perceived gender. To the extent that our perception of artist gender is similar to editors' perception of artist gender, ours is probably the relevant measure.

In addition to artist gender, we also observe artists' 2016 streams on Spotify, by country, as well as artist genre, according to Allmusic.com. Finally, we observe numerous song characteristics, including beats per minute, key signature, whether major or minor, and seven other characteristics of the songs measured on 100 point scales (danceability, valence, energy, acousticness, instrumentalness, liveness, and speechiness).

Excluding songs with artists of unknown gender among the first two, and including only songs with for which we have all relevant song characteristics, gives us an analysis sample including 6,650 songs. The songs in our analysis sample account for 32.2 billion streams, or for 97.4 percent of the total observable Spotify streams (the daily country top 200) for the

[8]Spotify makes these daily streaming data available at https://spotifycharts.com/regional.

period May 1 - December 31, 2017.

We then match the sample above with the songs appearing in our third dataset, which consists of the songs on the four most-followed editorial global playlists during most of 2017, Today's Top Hits, RapCaviar, ¡Viva Latino!, and Baila Reggaeton, three additional highly followed curated lists (Mint, Rock This, and Are & B).[9] Editorial lists differ in their genre focus and the number of followers (as of December, 2017): Today's Top Hits ("50 best from the world of music," 18.5 million followers), RapCaviar (rap, 8.5), ¡Viva Latino! ("top Latin hits", 6.9), Baila Reggaeton ("Latin Urban hits," 6.3), mint ("the world's top dance hits, 4.7), Rock This ("50 hottest rock songs," 4.0), Are & Be ("hottest R&B tunes," 3.9).[10] We observe the songs on these lists between May and December of 2017 as well as hand-collected gender information on the artists for each song.

Our fourth dataset covers the New Music Friday playlists, and our process for creating this sample is different. We started with all of the songs on the New Music Friday playlists in each of the 26 countries throughout 2017. Because we found impacts of New Music Friday list inclusion to be significant only for the top 20 songs (Aguiar and Waldfogel, 2018), we restricted attention to those songs in the weekly top 20. This is a total of 7,460 songs and 4,881 artists (as either the first or second artist on the songs) across these countries.

Because of the large number of New Music Friday artists, we used a three-pronged approach to determining gender for these artists. First, we searched the artist name on Musicbrainz, which provides genders for many individual artists. Second, we checked the artist names against the first name database. If we lacked a Musicbrainz match and the first name was attached to a gender with at least 90 percent probability, we assigned that gender. We found the remaining artists by hand. We discard songs when the genders of either of the first two listed artists is unknown.

Given our measures of gender for first and second-listed artists and bands in both the top artist and New Music Friday samples, we create four different binary measures of gender for a song, based on information on the first two listed artists or bands. In these measures, the

[9] Not all playlist songs match the 6,650 song top artist sample, for two reasons. First, some songs on the playlist between May and the end of the year first appeared in the streaming data prior to May and are therefore excluded from the 6,650. Second, some songs on the global playlists are by artists who do not have enough streams to appear in the top artist sample.

[10] For a playlist description see, for example, https://open.spotify.com/user/spotify/playlist/37i9dQZF1DWY7IeIP1cdjF.

song is deemed female if both artists are female ("all female"), if one of the two artists is female and the other is mixed ("female or mixed"), if the first listed artist is female, regardless of the second ("first artist female"), or if either of the two artists is female ("either female").

4 Results on female shares

4.1 What share of songs entering Spotify are by women?

The Everynoise data on songs entering Spotify include 1,833,874 songs entering Spotify between late in 2015 and early in 2018.[11] These songs have 765,619 distinct first artists associated with them, far too many for ascertaining gender by hand. Some of the artist names are bands, others are performers' names. We divide the artist names into elements separated by spaces, such that the first string is potentially an artist's first name. We then match this first element to databases on the gender frequency of names to get an estimate of the share of songs performed by female artists. This of course works only for solo artists (e.g. "Pete Yorn").

We are able to match an artist's name to gender for 558,528 of the 1,833,874 underlying songs, or just over 30 percent, in Table 2. Of these, roughly one in five of the songs overall has artists with female names. The share varies across the years, between 23.2 percent in (partial) 2015 and 20.8 percent in (partial) 2018 and 21.4-21.6 percent during the full years for 2016 and 2017. In other words, about 21.5 percent of the matched songs entering Spotify during 2016-2017 have artist names that are characteristically female.

Before going any further, it is worth noting the two facts identified thus far: about 38 percent of "musicians, singers, and related workers" are women, and just over a fifth of songs entering Spotify have female artists. These facts suggest that the female share of available music likely plays a role in determining the female share of songs succeeding on the platform.

[11] See http://everynoise.com/spotify_new_releases.html.

4.2 What share of major streaming songs are female?

Table 3 reports the female shares of songs and streams in the 6650 songs in the top artist sample. The female shares of songs vary substantially, between 13.0 percent for the most restrictive ("all female") measure and 22.0 percent for the most liberal "either female" approach. The respective female shares of streaming in this sample run from 12.4 percent and 31.8 percent.[12]

4.3 What share of playlists are occupied by female artists?

4.3.1 Global editorial lists

Pelly (2018) notes the low share of female artists on major playlists, and we see similar evidence here. Table 4 reports the female shares of songs - using the "first artist" approach - for each of the global playlists in the study. The table is based on all of the playlist songs in the data and not just those matched with the top artist sample. As Table 4 shows, the female shares are 12 percent on Viva Latino, 7.1 percent on Baila Reggaeton, and 6.3 on Mint. Shares are even lower on RapCaviar (6.1 percent) and Rock This (4.5 percent). While the female shares on these lists could reflect gender bias, they may also arise because of differential participation in the genres of music that the lists cover.

The female shares on Are & Be and Today's Top Hits are substantially higher, however. Of the 117 songs we observe on Are & Be, 23.1 percent are by female artists, and of the 221 songs that we observe on Today's Top Hits, 29.9 percent are by female artists. At Today's Top Hits, this is roughly a third greater than the female share of available or top songs.

4.3.2 New Music Friday lists

Figure 1 shows the female share of the artists for songs in the top 20 of those on the New Music Friday lists, by country, and across the four female measures. While levels vary across

[12] We get broad corroboration of our calculations from Everynoise, which provides a daily calculation of the share of Spotify streams for female artists. On August 10, 2018, the site reported that 21.3% of streams were "female or mixed-gender artists." See http://everynoise.com/gender_tldr.html. The female share fluctuates a bit but appears fairly stable: on August 23, 2018, it was 21.8 percent; and on September 14, 2018, it was 21.0.

female measures, the ordering of countries is similar. Using the "first artist" approach, for example, the female share runs from about 20 percent in Poland and Denmark to above 30 percent in Hong Kong, the Philippines, Singapore, and Malaysia.

4.3.3 Raw demographic shares and bias

The raw female shares of successful songs and playlist entries are, explicitly or implicitly, being interpreted as evidence of bias. This approach makes some intuitive sense, in that it is surprising that women are half as common as performers than as customers.[13] But the approach of inferring bias from a comparison of a group's share among successful artists and its share in the general population has limits. For example, African Americans make up 13.4 percent of US population, while they make up about 31.5 percent of the performers of songs on the Billboard Hot 100 in 2012-2017, as Figure 2 shows.[14] It is not clear that one could reasonably infer that non-blacks are the victims of bias in the music industry.

5 Testing for Bias

We have two broad ways to test for bias in the composition of playlists. The first broad approach conditions on observables and asks whether female-driven songs are differently likely to be included - or ranked - on playlists. The second approach examines the streaming success of playlisted songs, inferring that songs that perform better, conditional on the New Music Friday playlist rank, have faced bias.

5.1 Conditioning on observables: global editorial lists

A determination of bias requires a point of reference. For example, one might infer bias based on a comparison of the female stream share with the female song share. We can generalize this idea to control for a host of factors potentially predictive of songs being on a playlist rather than just comparing a female share of successful songs to a known benchmark

[13] According to everynoise.com, women made up nearly (44.7 percent) of Spotify listeners on November 5, 2018. See `http://everynoise.com/gender_tldr.html`.

[14] See `https://www.census.gov/quickfacts/fact/table/US/PST045217` and the Billboard Yearend lists, such as `https://www.billboard.com/charts/year-end/2017/hot-100-songs`. This is the sampling frame used by Smith et al. (2018).

such as the female share of the population. What can we use as control variables? We observe previous-year (2016) Spotify streams for each artist whose songs are in the top artist sample. To the extent that past success is predictive of current-song appeal, it may be a useful predictor of inclusion on major playlists.

We explore this idea in Figure 3, which shows the relationship between past streams (on the horizontal axes) and the probability that a song appears on Today's Top Hits, separately for male and female-driven songs and separately for each of the four female measures. Except for the group of songs whose artists lack 2016 streams in the data, the majority of the 2016 log streams values lie between 10 and 20 (i.e. the past streams lie between about 22,000 and 485 million). Hence, the graphs in Figure 3 are based on the observations in this range. Using the two most restrictive female measures, the relationship between past streams and list inclusion is positive for male but not for female songs. Using the two less restrictive female measures, by contrast, the probability of list inclusion rises with past streams for both male and female-designated songs. Moreover, the probability of inclusion on Today's Top Hits is higher for female than for male songs, conditional on past streams.

We can model the relationships in Figure 3 via regressions controlling for a host of song characteristics. These include whether the song is of US-origin, the number of beats per minute, seven characteristics of the songs measured on 100 point scales (danceability, valence, energy, acousticness, instrumentalness, liveness, and speechiness), whether in a major or a minor key, and the song's key signature. Moreover, participation in music may differ by genre, so we include dummies for each of 20 genres as well. Our genre measure refers to the first artist's reported genre in Allmusic. Panel A of Table 5 reports the resulting gender coefficients from linear probability model regressions of playlist inclusion on the songs' gender measure and controls. These control variables include the first artist's total streams in 2016, an indicator variable for artists who lack past streams in 2016, song-specific variables, including whether the song is of US-origin, the number of beats per minute, seven characteristics of the songs measured on 100 point scales (danceability, valence, energy, acousticness, instrumentalness, liveness, and speechiness), whether the song is in a major or a minor key, the song's key signature, and the genre of the song. The table reports the results for each playlist and gender measure. Because our approach asks which songs among the top artist sample (and which first appear in the streaming data on or after May 1, 2017) appear on the major playlists, the regressions include only playlisted songs that are also in the top artists sample. This gives us 64 songs from Baila Reggaeton, 42 for ¡Viva Latino!, 44 for Mint, 19 for Are & Be, 164 for RapCaviar, 141 for Today's Top Hits, and 25 songs for Rock This. The small number of top artist sample songs on Mint and Rock This suggest caution in the interpretation of results for those lists.

Of the 28 coefficients on the female dummies in Panel A, most are statistically insignificant, seven are significantly negative, and three are significantly positive. The RapCaviar coefficient is negative for three of four gender measures. The Rock This measure is negative for all four, although the top artist dataset includes only 25 songs on Rock This. Finally, the two Today's Top Hits coefficients from the models with the more liberal female measure are positive and significant.

One of the difficulties in attempting to control for relevant observables is the arbitrariness in the choice of control variables. Belloni et al. (2014) develop a "double selection" machine-learning method for choosing control variables in linear regression models whose purpose is causal parameter inference. The procedure is as follows. Given an outcome y, an explanatory variable of interest x, and a candidate list of controls Z, one runs LASSO regressions of y on Z, and x on Z, with LASSO's variable selection procedure selecting variables as they aid in reducing mean squared error. The union of the selected variables are then used as controls Z' in a regression of y on x and Z'.[15]

We assemble the candidate variables as follows. In addition to the variables in panel A of Table 5, we also include 19 genre dummies, 16 origin country dummies, as well as the full list of interactions (including squares) among all of the candidate variables. Panel B reports the coefficients on the female variable from the four double selection procedures associated with the eight playlists.

Results differ somewhat depending on the female measure employed in the machine learning approach in the bottom half of the table. Only three of the playlists provide any evidence consistent with bias: RapCaviar, Rock This, and Today's Top Hits. For three of four female measures, RapCaviar has a significantly negative female coefficient, indicating that songs with female involvement are less likely to appear on RapCaviar, conditional on other observables. Rock This, too, has a significantly negative coefficient using all four female measures, although it's important to recall that we have only 25 songs in the top artist

[15]See Ahrens et al. (2018b,a) for an implementation of this procedure in Stata.

sample that appear on Rock This. Finally, Today's Top Hits has a significant and positive coefficient for the two least restrictive female measures, reflecting a possible pro-female bias.

5.2 Conditioning on observables: New Music Friday lists

Artists' past streams are also potentially useful for predicting New Music Friday ranks among songs on the New Music Friday lists. Using the sample of songs in the top 20 of weekly New Music Friday ranks - and excluding (for the Figure) the songs whose artists lack past streams - most of the past log streams are between 12 and 20. Figure 4 shows how the smoothed average US New Music Friday rank varies with past streams by gender. Conditional on past streams, New Music Friday ranks are higher (worse) by about 1 for male artists.

Table 6 reports linear regressions of New Music Friday ranks on a progressively increasing set of variables. Column (1) controls for nothing, and the New Music Friday ranks are between 0.25 and 0.79 lower (better) for women. Column (2) adds country fixed effects, and the results do not change. Column (3) controls for past log streams (and an indicator for zero past streams) in country-specific ways, and the female coefficients now range between -0.65 and -0.79. The fourth column adds song characteristics as controls, such as whether the song is of US-origin, the number of beats per minute, seven characteristics of the songs measured on 100 point scales (described above), whether the song is in a major or a minor key, and the song's key signature. The coefficient range between -0.68 and -0.78. The final column approaches this question using the double selection ML approach, where the covariates include all the song-specific variables included in column (3), interacted with country dummies. The approach selects between 5 and 9 of 746 possible control variables for inclusion, and the resulting coefficients on the female variables run between -0.65 and -0.77 (with standard errors below 0.1). These regressions confirm that, among songs making the New Music Friday top 20, conditional on song and artist characteristics, the songs by women receive lower (better) ranks, by roughly 0.7 on average.

5.3 Outcome based bias tests

One inherent challenge with the conditioning on observables approach to measuring bias is a concern that the independent variable of interest - here, whether the song is by a female artist - is correlated with some important unobservable determinant of making the playlist. Addressing this requires a different method of measuring bias.

One such method is to measure bias based on streaming outcomes. Suppose an editor is choosing what rank to give a song on the New Music Friday list. If the editor is biased against some kinds of artists, then the editor will give those artists higher (worse) ranks than their songs warrant, in the sense of the songs' tendency to be popular and therefore streamed. A song facing bias against its creator will receive a worse rank than its underlying appeal should have suggested; as a result it will stream more, compared with songs with the same New Music Friday rank that don't face bias. Analogous tests for discrimination appear in Ayres and Waldfogel (1994), Smart and Waldfogel (1996), and Knowles et al. (2001).

A simple model makes the approach explicit. Song i has true quality q_i, which curators can ascertain with error ε_i. The curator's assessment of quality is thus $q_i' = q_i + \varepsilon_i$. Songs with higher assessed quality receive better ranks. That is, songs are ranked according to the order of q_i'. Because of their promotional value, better ranks raise streams proportionally by a_k, where $a_k > 1$, k is the New Music Friday rank, and $a_k \geq a_{k+1}$. In the absence of bias, songs with higher ranks are better and stream more, both because they are better and because they have received better ranks.

Now suppose there are two types of songs - by men and women - and that the editor is biased, so that the editor's assessment is $q_i' = q_i + \beta\delta^{female} + \varepsilon_i$. A bias against female artists means that $\beta < 0$. Then a song of a given quality level gets a worse rank if it's by a female artist. While the song now receives less ultimate streaming because it misses out on the higher promotional value of a better rank, it nevertheless has more underlying appeal than a male song receiving its New Music Friday rank. Hence, a symptom of bias against a female song would be that, given its rank, it would stream more than a similarly ranked male song.

Figure 5 implements this idea using a dummy for whether a song appears among the daily top 200 as the measure of streaming success, on the vertical axis. Songs' New Music Friday ranks appear along the horizontal axis, and the figure differentiates streaming success by artist gender. A few things are evident from the figure. First, songs with better New Music Friday ranks are more likely to appear in the top 200. Second, songs by women appear - at least sometimes - to have different streaming success at any given New Music Friday rank. For example, songs by female artists ranked 11-20 tend to have lower streaming success than songs by male artists, and this emerges for all four female measures.

We explore this more systematically with regressions, in Table 7. The first set of columns uses the top 20-ranked songs on New Music Friday playlists, with three separate dependent variables measuring streaming success. The first column's dependent variable is a binary measure of whether the song appears in the daily top 200. The dependent variable in the second column is an indicator for whether the song appears in the daily top 100. The remaining columns repeat the exercises using the top 10 songs on the weekly New Music Friday lists rather than the top 20.

Echoing the suggestion of Figure 5, we see relatively little evidence of bias in the top 10. Negative and significant coefficients, reflecting pro-female bias, appear for the two more restrictive gender measures and for the outcome variables indicating top 100 appearance but not for top 200 appearance. Evidence of bias is more systematic in the top 20 overall (driven by songs at 11-20). Coefficients are negative and significant for both streaming success measures, using all but the most liberal female gender measure. Overall, the New Music Friday rankings appear to incorporate a pro-female bias: rankings are statistically significantly better (by about two thirds of a rank), and - except with the most inclusive female measure - subsequent streaming performance is statistically significantly lower.

5.4 Magnitudes

Table 5 provides some fine-grained evidence of gender bias in global playlist inclusion, and it would be useful to measure the magnitude of effect of biases on the female streaming share. We can use the bias coefficients in Table 5 to determine how many songs by female artists should have made one of the playlists. To illustrate, consider the LASSO estimate using the "either female" measure for Today's Top Hits, 0.0144. Using the "either female" measure, 1,463 songs are female, and the associated female dummy implies that 21 too many female songs (or 0.0144 x 1,463) made Today's Top Hits than would have happened in the absence of bias. We can make this adjustment for each of the playlists and each of the female measures. Adding the number of female songs across lists, we find 30, 33, 76, and 121 in the actual data for the four successive gender measures. Eliminating the bias using the estimated female bias coefficients would change these numbers to 63, 64, 83, and 117, respectively.

We can roughly remove the impact of bias on the female shares of streaming by adding or subtracting female songs from the playlists, then assuming they would have received the additional streams caused by list inclusion. In Aguiar and Waldfogel (2018) we document that inclusion on one of the four largest global playlists (Today's Top Hits, RapCaviar, Viva Latino, or Baila Reggaeton) raises a song's streams by an average of 26.8 million streams. These lists have more followers than the other three, so these lists' effects probably exceed those for the others. Our adjustments will therefore be arguably be upper bounds.

Using the "all female" measure, the actual sample includes 30 playlisted songs in this group, while de-biasing would raise this to 63. Hence, the de-biased "all female" streaming total is the actual female streaming total, plus 33 additional playlisted songs times 26.8 million streams per song, while the corresponding male total is the male actual total, less 33 times 26.8 million. Table 8 reports the de-biased female shares for all four female measures along with the actual measure. Using the two most restrictive gender measures, de-biasing raises the female share of total streaming to nearly 15 percent. Using the two less restrictive measures the female share is fairly stable around 23 and 32 percent, respectively.

Despite the evidence of gender bias in Table 5, this bias accounts for relatively little of variation in gender shares. The actual and de-biased female shares in Table 8 are similarly low. Rather than bias on the platform, the gender shares arise from the differences in gender shares of songs on the platform.

6 Conclusion

The rise of Internet platforms has focused attention on their power and the possibility of biased exercise of that power. The music industry has, at the same time, faced scrutiny for the relatively small share of women among streaming music totals. Female success lags far behind male success, with women accounting for between a seventh and nearly a third of streaming, depending on the gender measure employed. Using the Spotify platform as a testing ground, it appears that most of the relatively low shares of female song streaming stems from relatively few female songs on the platform, rather than from bias on the platform. Evidence on bias in the platform's playlist decisions is mixed. After accounting for plausible determinants of inclusion on playlists, songs by female artists are substantially more likely to be included on Today's Top Hits and less likely to appear on RapCaviar. Finally, songs by female artists obtain lower (better) New Music Friday ranks than songs by male artists, after accounting for past streams and song characteristics. We find some evidence that these better rankings reflect bias. Conditional on New Music Friday rank, songs by female artists are less likely to achieve streaming success. While we find some evidence consistent with bias - both for and against female artists - this explains little of the low female share of streaming. We attribute that, instead, to the low female share of songs on the platform.

We do not interpret our ability to explain the low female share of successful songs as a justification of the low shares. One way to raise the female share of successful songs would be for playlists to feature those songs more prominently. In some respects this is already happening, for example with female artists facing a higher likelihood of appearing on Today's Top Hits or obtaining better ranks on the New Music Friday lists. But our results suggest that efforts to raise the female share of artists on the platform might be fruitful steps toward raising the female share of successful songs. It is not clear how much of the low share of female songs on the platform arises from lower female participation in music creation, as opposed to lower rates of bringing music to audiences by putting music on the platform. It is perhaps encouraging that women make up nearly 38 percent of musicians in the US, compared with roughly a fifth of songs entering Spotify. Efforts to get more music onto the Spotify platform might give rise to higher female shares of streaming. Finally, the issues addressed in this study are by no means specific to Spotify. If data were available, it would be desirable to ask this study's questions of other platforms.

References

Aguiar, L. and J. Waldfogel (2018): "Platforms, Promotion, and Product Discovery: Evidence from Spotify Playlists," *JRC Digital Economy Working Paper 2018-04*.

Ahrens, A., C. B. Hansen, and M. Schaffer (2018a): "LASSOPACK: Stata module for lasso, square-root lasso, elastic net, ridge, adaptive lasso estimation and cross-validation," *Boston College Department of Economics*.

——— (2018b): "PDSLASSO: Stata module for or post-selection and post-regularization OLS or IV estimation and inference," *Boston College Department of Economics*.

Ayres, I. and J. Waldfogel (1994): "A market test for race discrimination in bail setting," *Stanford Law Review*, 987–1047.

Belloni, A., V. Chernozhukov, and C. Hansen (2014): "High-dimensional methods and inference on structural and treatment effects," *Journal of Economic Perspectives*, 28, 29–50.

Edelman, B. (2011): "Bias in search results: Diagnosis and response," *Indian JL & Tech.*, 7, 16.

Ellis-Petersen, H. (2014): "Gender bias in the film industry: 75% of blockbuster crews are male," *The Guardian*, https://www.theguardian.com/film/2014/jul/22/gender-bias-film-industry-75-percent-male, July 22.

Flanagan, A. (2018): "Grammy President Neil Portnow To Step Down In 2019." *NPR.com*, https://www.npr.org/sections/therecord/2018/06/01/615889769/grammy-president-neil-portnow-to-step-down-in-201, June 1.

Hegewisch, A. and H. Hartmann (2014): "Occupational Segregation and the Gender Wage Gap: A Job Half Done." *Institute for Women's Policy Research (under contract with the U.S. Department of Labor)*, https://www.dol.gov/wb/resources/occupational_segregation_and_wage_gap.pdf, January.

Knowles, J., N. Persico, and P. Todd (2001): "Racial bias in motor vehicle searches: Theory and evidence," *Journal of Political Economy*, 109, 203–229.

Lambrecht, A. and C. E. Tucker (2018): "Algorithmic bias? An empirical study into apparent gender-based discrimination in the display of STEM career ads," .

Pelly, L. (2018): "Discover Weakly." *The Baffler*, https://thebaffler.com/latest/discover-weakly-pelly., June 4.

Smart, S. and J. Waldfogel (1996): "A citation-based test for discrimination at economics and finance journals," *National Bureau of Economic Research*.

Smith, S. L., M. Choueiti, and K. Pieper (2018): "Inclusion in the Recording Studio? Gender and Race/Ethnicity of Artists, Songwriters & Producers across 600 Popular Songs from 2012-2017," *Annenberg Inclusion Initiative*, http://assets.uscannenberg.org/docs/inclusion-in-the-recording-studio.pdf.

Zhu, F. and Q. Liu (2016): "Competing with complementors: An empirical look at amazon.com," .

A Figures and Tables

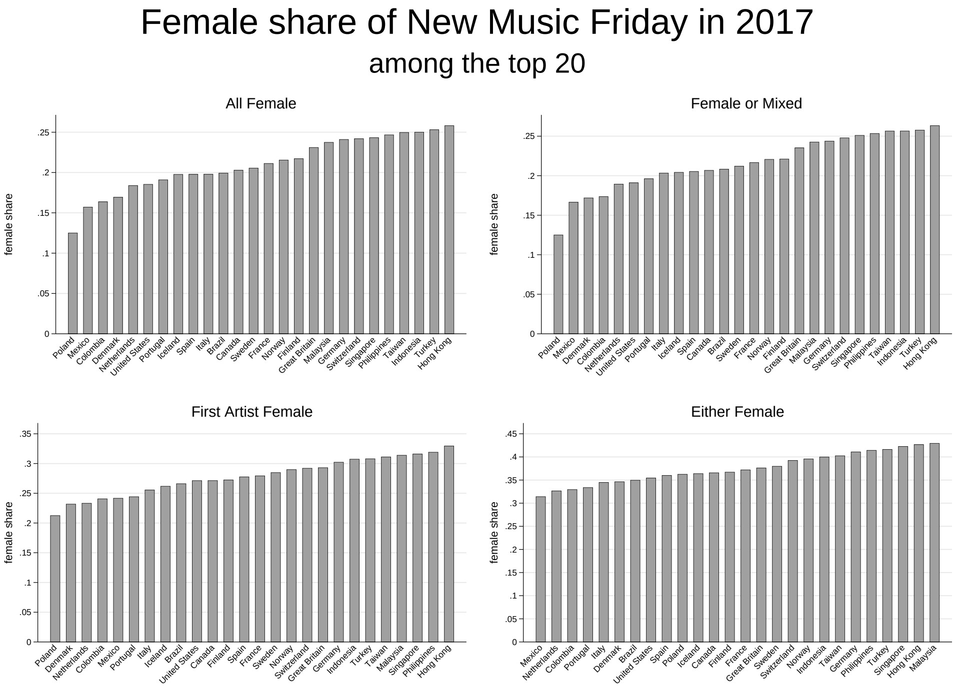


Figure 1: Female Share of New Music Friday, Top 20.

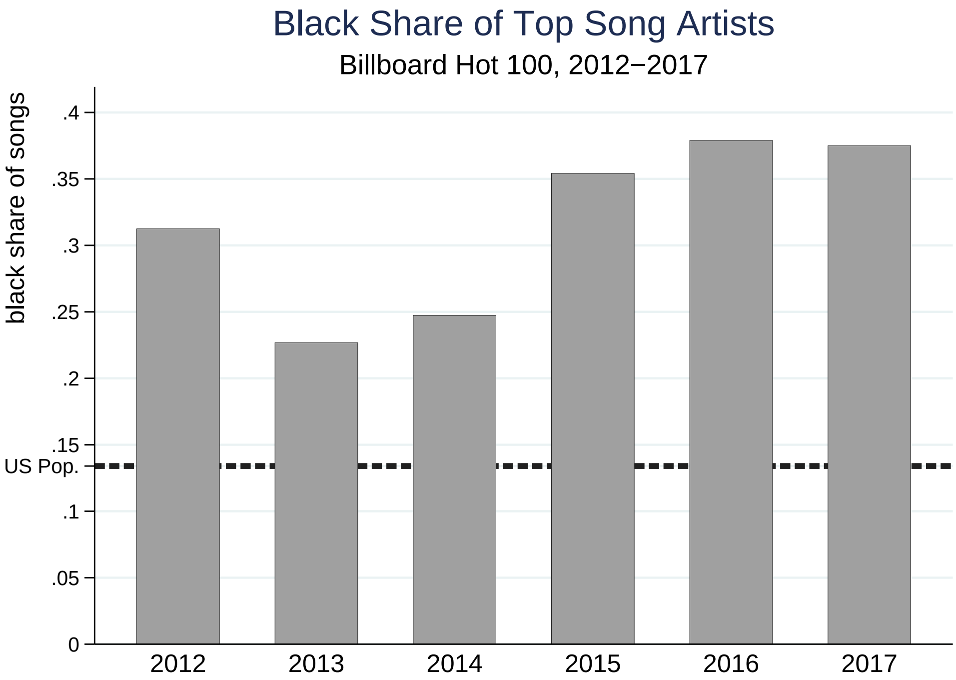


Figure 2: Black Share of Top Song Artists in US.

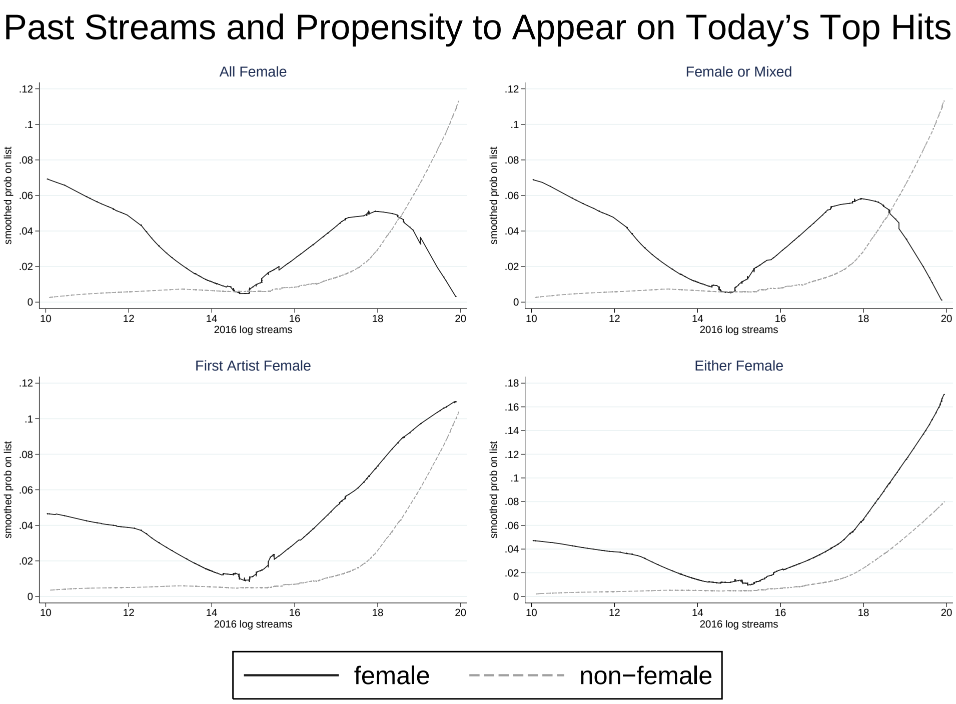


Figure 3: Past Streaming and Propensity to be on Today's Top Hits.

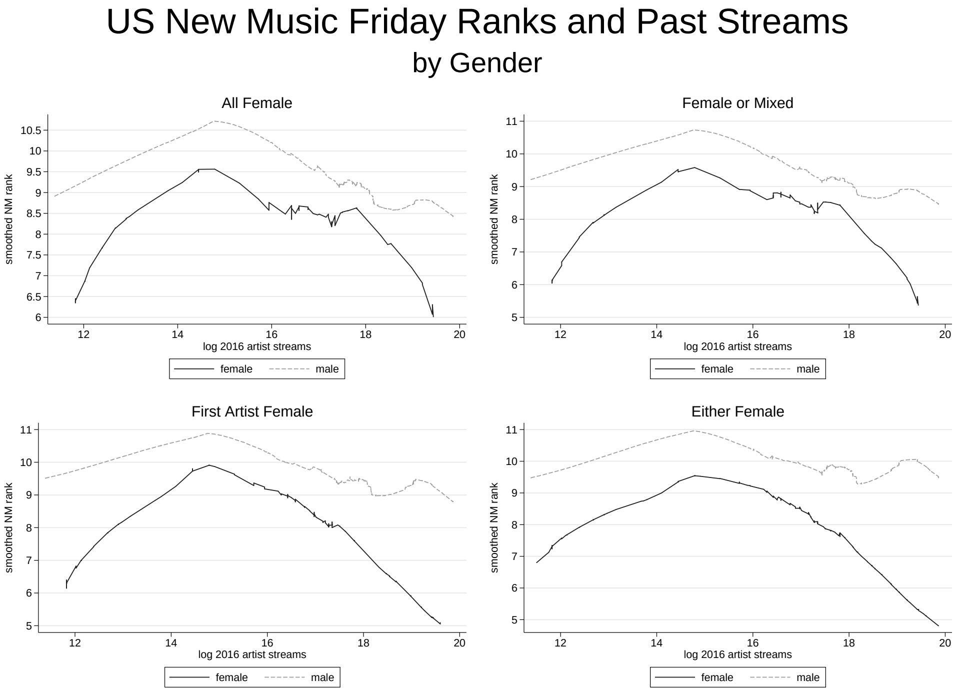


Figure 4: US New Music Friday Ranks and Past Streams by Gender.

Figure 5: Making the Top 200, by New Music Friday Rank and Gender.

Table 1: Demographic shares of arts, design, entertainment, sports, and media occupations.[†]

Occupation	**Employment (in thousands)**	**Women**	**White**	**Black**	**Asian**	**Hispanic or Latin**
Arts, design, entertainment, sports, and media occupations	3,246	46.9	83.9	7.3	5.3	10.2
Artists and related workers	236	51.9	89.6	3.5	4.5	10.8
Designers	908	52.8	84.2	5.1	7.6	12
Actors	40					
Producers and directors	180	29.8	87.8	5.2	4.6	8.9
Athletes, coaches, umpires, and related workers	345	36.4	88.2	5	3	8.5
Dancers and choreographers	21					
Musicians, singers, and related workers	**188**	**37.7**	**75.6**	**17.3**	**4.3**	**6.7**
Entertainers and performers, sports and related workers, all other	67	48.1	74.1	9.2	9.9	7.5
Announcers	50	23.1	77.5	12	7.1	8.5
News analysts, reporters and correspondents	84	55.3	73	13.4	7.9	16.4
Public relations specialists	118	60.9	87.9	8.3	2.6	5.7
Editors	174	54.1	90.9	3.8	2.9	4.1
Technical writers	76	60.2	84.6	6.7	5.6	1.2
Writers and authors	226	60.9	86.3	6.3	3.2	4.6
Miscellaneous media and communication workers	112	64.2	70.6	9.5	14.3	32.9
Broadcast and sound engineering technicians and radio operators	119	14.4	78.2	14	3.7	16.8
Photographers	214	46	84.4	9.1	3	8.7
Television, video, and motion picture camera operators and editors	82	12.2	86.1	7.2	1.9	7.6
Media and communication equipment workers, all other	5					

[†] Employed persons by detailed occupation, sex, race, and Hispanic or Latino ethnicity, Arts, design, entertainment, sports, and media occupations, Bureau of Labor Statistics, January 19, 2018, https://www.bls.gov/cps/cpsaat11.htm.

Table 2: How female are the songs entering Spotify?[†]

Year	% Male	Matches	Total
2015	79.36%	19,463	63,377
2016	80.68%	179,426	567,692
2017	80.92%	281,458	934,261
2018	81.59%	78,181	268,544

[†] Source: author's calculation from Everynoise.

Table 3: Gender Shares of Songs and Streams.[†]

	All Female	Female or Mixed	First Artist Female	Either Female	Number of Songs
New Music Friday	21.3%	21.9%	28.0%	37.5%	18489
Today's Top Hits	16.3%	17.7%	29.8%	41.1%	141
¡Viva Latino!	2.4%	2.4%	21.4%	40.5%	42
Baila Reggaeton	3.1%	3.1%	10.9%	23.4%	64
RapCaviar	1.2%	1.8%	7.9%	11.6%	164
Mint	0.0%	0.0%	4.5%	18.2%	44
Are & Be	5.3%	5.3%	10.5%	15.8%	19
Rock This	4.0%	4.0%	4.0%	4.0%	25
Top Artists Sample	13.0%	13.2%	16.4%	22.0%	6650
Top Artists Sample - Streams	12.4%	12.7%	22.5%	31.8%	6650

[†] For each of the global playlists, the table includes songs that appeared in the Top Artists sample as explained in the text.

Table 4: How female are the global playlists?†

Playlist	Female Share	Number of Songs
Baila Reggaeton	7.1%	140
Mint	6.3%	810
Are & Be	23.1%	117
RapCaviar	6.1%	393
Rock This	4.5%	176
Today's Top Hits	29.9%	221
¡Viva Latino!	12.0%	108

† Songs are defined as female if the first artist is a female.

Table 5: Global List Inclusion (from May on).†

	(BR) Coef./s.e.	(MT) Coef./s.e.	(RB) Coef./s.e.	(RC) Coef./s.e.	(RT) Coef./s.e.	(TT) Coef./s.e.	(VL) Coef./s.e.
A. Linear probability model							
All Female	-0.0039	-0.0035	-0.0017	-0.0116**	-0.0084***	0.0015	-0.0042
	(0.004)	(0.003)	(0.002)	(0.006)	(0.002)	(0.005)	(0.003)
Female or Mixed	-0.0047	-0.0039	-0.0018	-0.0104*	-0.0084***	0.0036	-0.0046
	(0.004)	(0.003)	(0.002)	(0.005)	(0.002)	(0.005)	(0.003)
First Artist Female	-0.0011	-0.0026	-0.0023	-0.0066	-0.0093***	0.0145***	0.0038
	(0.003)	(0.003)	(0.002)	(0.005)	(0.002)	(0.005)	(0.003)
Either Female	0.0009	-0.0031	-0.0024	-0.0102**	-0.0079***	0.0144***	0.0070***
	(0.003)	(0.002)	(0.002)	(0.004)	(0.002)	(0.004)	(0.002)
No. of Obs.	6650	6650	6650	6650	6650	6650	6650
B. LASSO double selection							
All Female	-0.0050	-0.0040	-0.0013	-0.0122**	-0.0075***	-0.0032	-0.0047
	(0.004)	(0.003)	(0.002)	(0.005)	(0.002)	(0.005)	(0.003)
Female or Mixed	-0.0056	-0.0043	-0.0012	-0.0109**	-0.0076***	-0.0009	-0.0051*
	(0.004)	(0.003)	(0.002)	(0.005)	(0.002)	(0.005)	(0.003)
First Artist Female	-0.0012	-0.0022	-0.0020	-0.0058	-0.0088***	0.0102**	0.0030
	(0.003)	(0.003)	(0.002)	(0.005)	(0.002)	(0.005)	(0.003)
Either Female	0.0011	-0.0030	-0.0016	-0.0075*	-0.0076***	0.0144***	0.0069***
	(0.003)	(0.002)	(0.002)	(0.004)	(0.002)	(0.004)	(0.002)
No. of Obs.	6650	6650	6650	6650	6650	6650	6650

† Specifications in Panel A include the first artist's total streams in 2016 as well as an indicator variable for artists who lack past streams in 2016. Song-specific variables are also included, including whether the song is US-origin, the number of beats per minute, seven characteristics of the songs measured on 100 point scales (danceability, valence, energy, acousticness, instrumentalness, liveness, and speechiness), whether the song is in a major or a minor key, the song's key signature and the genre of the song. Panel B includes control variables selected through the LASSO prcedure as described in the text.

* Significant at the 10% level.

** Significant at the 5% level.

*** Significant at the 1% level.

Table 6: New Music Friday Playlist Rankings (among Top 20). †

	(1) Coef./s.e.	(2) Coef./s.e.	(3) Coef./s.e.	(4) Coef./s.e.	(lasso) Coef./s.e.
All Female	-0.2516**	-0.2529**	-0.6489***	-0.6844***	-0.6524***
	(0.105)	(0.105)	(0.100)	(0.102)	(0.101)
Female or Mixed	-0.3568***	-0.3587***	-0.7181***	-0.7479***	-0.7179***
	(0.104)	(0.104)	(0.099)	(0.101)	(0.100)
First Artist Female	-0.6332***	-0.6362***	-0.7874***	-0.7845***	-0.7706***
	(0.095)	(0.096)	(0.091)	(0.093)	(0.092)
Either Female	-0.7857***	-0.7894***	-0.7524***	-0.7197***	-0.7327***
	(0.088)	(0.088)	(0.084)	(0.085)	(0.085)
Country Fixed Effects	✗	✓	✓	✓	✓
Past Streams × Country FE	✗	✗	✓	✓	✓
Song Characteristics	✗	✗	✗	✓	✓
No. of Obs.	18233	18233	18233	18233	18233

† The dependent variable is the ranking position on the New Music Friday playlist. Each row corresponds to a distinct regression.
* Significant at the 10% level.
** Significant at the 5% level.
*** Significant at the 1% level.

Table 7: Streams Conditional on Rank and Gender. †

	Top 20 New Music		Top 10 New Music	
	(Top200) Coef./s.e.	(Top100) Coef./s.e.	(Top200) Coef./s.e.	(Top100) Coef./s.e.
All Female	-0.0369***	-0.0349***	-0.0031	-0.0314***
	(0.008)	(0.007)	(0.012)	(0.011)
Female or Mixed	-0.0385***	-0.0359***	-0.0068	-0.0339***
	(0.007)	(0.006)	(0.012)	(0.011)
First Artist Female	-0.0155**	-0.0137**	0.0191*	-0.0011
	(0.007)	(0.006)	(0.011)	(0.010)
Either Female	-0.0009	-0.0025	0.0289***	0.0144
	(0.006)	(0.006)	(0.010)	(0.009)
No. of Obs.	18233	18233	9193	9193

† All specifications include rank dummies. Each row corresponds to a distinct regression.
* Significant at the 10% level.
** Significant at the 5% level.
*** Significant at the 1% level.

Table 8: Actual and de-biased female streaming shares in the top artist sample. †

	All Female	Female or Mixed	First Artist Female	Either Female
Actual female songs	13.0%	13.2%	16.4%	22.0%
Actual female streams	12.4%	12.7%	22.5%	31.8%
De-biased female streams	14.7%	14.9%	22.9%	31.6%

† De-biased female streams are actual streams, adjusted for the number of female songs that would have been on the study's playlists absent the bias documented in Table 5. Songs added to playlists are assumed to receive an additional 26.8 million streams per song, based on evidence in Aguiar and Waldfogel (2018).

JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.