Streaming Services and the Homogenization of Music Consumption*

George Knox†
Hannes Datta‡

This Version: February 23, 2020

* The authors are grateful for comments from seminar participants at Columbia University, Duke University’s Fuqua School of Business, Eindhoven University of Technology, Goethe University Frankfurt, University of Hamburg, the University of Minnesota, the University of Rochester, and the University of Zurich. The authors also thankfully acknowledge valuable comments and discussions with Bart Bronnenberg, Hana Choi, Dean Eckles, Eric Johnson, Carl Mela, Rachel Meng, Chris Moorman, Thomas Otter, and Raphael Thomadsen. The authors further acknowledge financial support from the Netherlands Organization for Scientific Research (NWO 451-17-028) and from MSI (MSI 4-1854).
† Associate professor of marketing at Tilburg University, g.knox@tilburguniversity.edu
‡ Associate professor of marketing at Tilburg University, h.datta@tilburguniversity.edu
STREAMING SERVICES AND
THE HOMOGENIZATION OF MUSIC CONSUMPTION

Given massive assortment and limited consumer knowledge, policymakers and the media are concerned that platforms may homogenize tastes by pushing consumers to the same content. To study this issue, we analyze music consumption for pairs of consumers around the time they adopt Spotify, a major music streaming platform. We find that Spotify adoption indeed makes listening behavior more similar across consumers, and we interpret this by decomposing it into two effects. First, Spotify expands the size of the consumption set, increasing expected listening similarity even in the absence of any platform-directed coordination. Second, accounting for consumption set expansion, adoption of Spotify decreases similarity in content consumption. These effects increase over time and are stronger for heavy users. Rather than making large groups of consumers listen to the same content, this paper, therefore, suggests streaming services expand and individualize listening behavior, something consumers are likely to find valuable. We discuss implications for musicians, labels, and marketers.

Keywords: streaming platforms, music consumption, variety, difference-in-difference
Digitization has lowered the costs of distributing products in cultural product markets (Waldfogel 2017). These markets now collect content across dispersed suppliers and offer large assortments to consumers. Online streaming accounted for 47% of global revenue in the music industry in 2018 (Perez 2019) and is rapidly growing in video, books, and games. The largest platform, Spotify, has 100 million paid subscribers, across 79 countries, choosing from a catalog of over 35 million songs (Spotify 2019).

How do consumers choose from such massive assortments? First, cognitive limits severely restrict consumer’s awareness and knowledge of the vast majority of Spotify’s catalog. Second, even if consumers were aware of a particular song and liked it, unassisted consumption would require retrieving it from memory and ranking it sufficiently high in terms of salience. Consumers are constrained by the mental and time costs of searching, memory retrieval, and ranking of many alternatives. Platforms can shape the consumer search process by choosing which content appears at the top of the screen, deciding how content is curated (e.g., bundled into playlists), and employing different technologies to learn about their users’ tastes to recommend content to them (Goldfarb and Tucker 2017; Ursu 2018). Demand depends not only on consumer preferences but also on preferences for time or convenience in accessing content.

A major instrument for platforms to lower search costs is the use of playlists, which account for 31% of all listening (Spotify 2017). Playlists come in many forms, but there are at least two fundamentally different classes of them (Aguiar and Waldfogel 2018a). First, platforms rely on recommender systems to create personalized playlists, like Discover Weekly and Release Radar in the case of Spotify. The versioning of these playlists is often based on behavioral targeting variables and past consumption. Second, platforms use playlists that are the same for everyone. The top five most popular playlists—Today’s Top Hits (with 24 million followers),
Global Top 50 (14), RapCaviar (12), Baila Reggaeton (10) and Viva Latino (10)—are globally similar, not personalized, and effectively similar to radio, augmented with out-of-sequence access functionality. They account for up to 20% of streams (Economist 2018) and have a large and significant impact on consumption (Aguiar and Waldfogel 2018a).

Thus, while streaming platforms combine technology and individual listening histories to personalize recommendations from their massive catalog, they also treat large groups of consumers with the same bundles of variety. The net effect of these opposing forces is unknown but has important implications for consumers, producers, and marketers. Policymakers and the media worry that Spotify may harm consumers by homogenizing tastes: “listeners who rely on Spotify to introduce them to new music are served up a relatively small number of artists – the ones that everyone else is listening to and the algorithm is monitoring” (Guardian 2019). There is also concern that platforms create unequal, winner-take-all outcomes for producers, e.g., musicians and labels on the platform (Aguiar and Waldfogel 2018a; Krueger 2019). According to the Economist, “Spotify’s most obvious power is to make stars via its playlists and recommendation algorithms, much as radio DJs used routinely to do with simple airplay” (2018).

Lastly, companies are increasingly shifting their marketing budgets to streaming platforms because they allow them to use consumption data to target consumers more efficiently (Bruell 2019). However, if platforms cause users to consume more similar bundles of variety, their ability to target them will be limited.

---

1 Hybrid forms exist too, e.g., the playlist New Music Friday is not personalized, but different for each country.
2 In Web Appendix A, we show that the treatment effect of similarity and concentration metrics like the HHI are strongly and positively related.
3 Spotify’s recently developed Ad Studio is designed to allow advertisers to choose their audience of the free ad-supported service based on “age, gender, location, activity [e.g., playlists for running, cooking], and even music taste.”
Using consumer-level data on music listening before and after the adoption of Spotify, we first investigate whether streaming platforms make music consumption more heterogeneous or, conversely, homogenous. Consumers have idiosyncratic tastes for music consumption subject to search frictions. Hence, a sensible research strategy of studying the overall effects of platform-directed search is to focus on whether Spotify adoption leads to consumers listening to more (or less) similar music.

Secondly, we decompose the overall Spotify adoption effect on consumption similarity into two components (Hosanagar et al. 2014). Streaming services entice consumers to expand their consumption set by offering variety at zero marginal cost (Datta, Knox, and Bronnenberg 2018). This increase in the size of the consumption set makes overlap more likely absent any platform-directed search. The second component refers to the similarity from the content of the consumption set, conditional on its size. Unlike the first component, the content of the consumption set may be influenced by either (algorithmically) personalized or editorially curated same-for-everyone playlists or other platform decisions on how to promote content. The decomposition has implications for consumer welfare. If Spotify increases similarity because its users consume more varieties overall, rather than switching among varieties, concerns about potential consumer harm due to platform-directed search seem unwarranted.

We provide evidence on the direction and magnitude of these two components. We develop a closed-form expression for (expected) similarity conditional on consumption set size, assuming no coordination in consumption. We use this variable to offset observed consumption similarity, thereby isolating the effect of Spotify adoption on the second component, i.e., the content of the consumption set. Because the total impact of adoption on similarity is the sum of
the two components, we can estimate two equations—one for the combined effect, and one for
the content component—to quantify and decompose the impact of Spotify adoption on similarity.

Lastly, we investigate how similarity in music consumption depends on the duration of
being a Spotify user. When a consumer first joins the service, the platform does not have any
consumption history; hence, its ability to present new customers personalized bundles of variety
is limited. As consumption data builds up over time, however, it can tailor its recommendations.
The change in these effects over time—e.g., whether user consumption similarity remains stable
or steadily decreases over time—is informative about how important experience (operationalized
as time-since-adoption of Spotify) is to personalization.

Although user similarity is well-defined for pairs of users, measuring the effect of Spotify
adoption on listener similarity is complicated by two confounding factors. Consumers may listen
to similar music before joining a streaming platform. Given our rich time-series data of users’
listening histories, we use fixed effects for each pair of users to control for pre-existing
consumption similarity. Second, content varies in overall popularity; in assessing any Spotify-
driven effects, co-consumption of popular content is less informative about similarity than co-
consumption of unpopular content. We follow the information retrieval literature (e.g., Elkan
2005) to construct our similarity metrics by weighting content inversely by its popularity
(measured across all platforms, not only Spotify).

Our data come from a music recommendation service that records digital listening on
many music platforms (i.e., also non-streaming platforms such as iTunes or Windows Media
Player, see Datta, Knox, and Bronnenberg 2018), allowing us to measure similarity before and
after consumers start using Spotify. We consider music consumption in terms of songs and
artists⁴, measuring similarity along both extensive and intensive margins. We collect listening histories and platform data from 2,311 users (561 of whom adopt Spotify) over 17 four-week periods, yielding 33 million pair-period observations from 2.7 million listener pairs.

First, we find that consumption similarity increases after Spotify adoption. However, the decomposition shows that this effect occurs because of the size effect, users expanding their consumption set, rather than users switching across varieties, content effect. After controlling for size, Spotify actually decreases similarity in terms of listening content. These results are largely robust to measuring similarity using songs and artists, using different similarity metrics, measuring similarity along extensive and intensive margins, and using different weights in accounting for the differential popularity of content. In line with the notion that time on the platform allows Spotify to personalize its recommendations better, this content effect grows more negative over time and is larger for heavy (vs. light) Spotify users.

This paper is organized as follows. First, we discuss the relevant literature. Second, we discuss our empirical measures, the data, and methodology. Finally, we present the empirical findings and discuss their implications.

**LITERATURE**

Our work is broadly related to research on the economic effects of digitization (Greenstein, Lerner, and Stern 2010) and on how consumers make choices in vast assortments (Goldfarb and Tucker 2017).

Digitization has lowered the costs of distribution and increased the variety of products available, opening up the “long tail” of rarely consumed content (Anderson 2006). Uncertainty in

---

⁴ We do not measure consumption in terms of genre because genres are less well-defined than songs and artists.
product quality means that some products with low expected profitability ex-ante would not have been distributed before digitization, but turn out to be highly desired by consumers (Aguiar and Waldfogel 2018b); several constructed quality indices for newly released music, movies, books, and television have accordingly risen since around 2000 (Waldfogel 2017).

However, large product assortments are more challenging for consumers to sift through to find the most preferred products. Firms can influence this search process in several ways, mostly through selectively lowering search costs or improving the efficiency of search (Yang 2013). For instance, sellers can influence the search process by choosing which attributes are salient. Using an experiment with two fictitious wine sellers online, Lynch and Ariely (2000) show that subjects focus on price when it is available on the first page, but they focus on other attributes when they need to click further to learn about price.

Online sellers can gather data on their users and deploy technology such as recommender systems to help consumers interact with large-scale assortments. There is a vast literature on algorithm design of such systems (Adomavicius and Tuzhilin 2005), and considerable research on optimal personalization in marketing (see Chung, Wedel and Rust 2016 for a review), however little is known about how much these systems empirically affect consumers (Sorenson 2017).

Krueger (2019) argues contrary to the long-tail hypothesis that music preferences are “highly skewed toward a relatively small group of musicians” in part because “systems that use Big Data to help us discover new music are also likely to reinforce network effects, unless there is a surge in demand for systems that recommend songs that are unpopular and likely to stay that way.” A few studies in the information systems literature examine the effects of recommender systems on consumer preferences and inequality in terms of the distribution of demand.
Adomavicius et al. (2013; 2018) show in several experiments that consumer preferences are malleable and strongly affected by recommendations. In one experiment, subjects were provided with randomly generated song recommendations and had to give their willingness-to-pay judgment. Even after sampling each song for 30 seconds, reducing uncertainty about the fit of the song, subjects gave higher valuations to songs that had more randomly assigned stars. Fleder and Hosanagar (2009) and Oestreicher-Singer and Sundararajan (2012) study the potential for recommender systems to concentrate consumption in the head or tail of the distribution. For example, the former study uses analytical models and simulation to show that recommender systems can end up suggesting the most popular products instead of promoting products in the “long tail” as commonly thought. One reason is that recommender systems are evaluated based on predictive ability (i.e., did the user choose a product that would have been recommended) rather than causal impact (i.e., how the user would have behaved in the absence of a recommender system). If many people buy Harry Potter, the recommender system that recommends Harry Potter to everyone will do well in terms of predicting out-of-sample choices. Secondly, collaborative filtering requires that products are bought at least once to be recommended; practitioners may use even higher thresholds before content is recommended.

The study closest in aim to ours is Hosanagar et al. (2014), who investigate the effects of a recommender plugin on iTunes on consumption similarity. They find that adopters increased their similarity among other adopters both because they purchased more songs (the “volume” effect) and because they shifted their listening towards more popular songs (the “product-mix” effect). We differ in several respects from their study. We investigate platform, rather than recommender system, adoption within a streaming, rather than ownership, business model, and we measure similarity in terms of consumption across many music listening platforms, not only
iTunes. Secondly, our decomposition methods differ: we measure changes in users’ consumption set sizes at the individual, rather than group, level, and we develop an analytical expression based on probability limits rather than a simulation approach. Our long panel data and staggered adoption times allow us to study effects over multiple periods so that we can assess the speed with which platforms shape co-consumption. Personalization of content may take some time as algorithms learn about user preferences, and consequently, we expect adoption effects to differ over time.

**MEASURES**

*Measuring Similarity*

Our objective is to estimate the effect of Spotify adoption on pairwise consumption similarity. We model listening behavior over consecutive 4-week periods, but we suppress time subscripts until our model section.

We measure the similarity between the listening histories of two consumers as follows. Suppose there are $N$ unique units (e.g., songs, or artists) in the union of listening histories across all consumers $i = [1, ..., I]$. The $(I \times N)$ matrix $X$ records consumption across consumers and songs. The elements of $X$ may be extensive (binary) measures of consumption—e.g., does user $i$ listen to song $k?$—or intensive measures such as the log count of the number of times a song was heard (i.e., log playcount). Measuring the consumption similarity between any two users $i$ and $j$ amounts to comparing row-vectors $X_i$ and $X_j$. Standard measures of similarity such as the Euclidean distance between the consumption-behavior of two consumers suffer from the drawback that two users who listen to the same set of songs but differ in their overall amount of listening will be dissimilar. Instead, for our purposes, if consumers listen to the same collection
of songs, we deem these consumers to be alike. Therefore, we turn to a commonly used measure called Cosine Similarity, which is equal to the inner product of the consumption vectors from consumers $i$ and $j$, divided by the product of the lengths of the two vectors (Murphy 2012), i.e.,

$$\text{Cosine Similarity}_{ij} = \frac{X_i X'_j}{\|X_i\| \|X'_j\|},$$

(1)

where $\|X_i\|$ is the Euclidean or $L^2$-norm of $X_i$. This measure derives its name from being equal to the cosine of the angle between the two vectors. It is bounded between 0 and 1. The more users $i$ and $j$ listen to the same songs, the more their vectors point in the same direction, making the angle smaller and the cosine of that angle closer to one. If there is no consumption overlap between users ($X_i X'_j = 0$), the cosine similarity is zero.

If the elements of $X$ are binary, we can simplify this expression to

$$\text{Cosine Similarity}_{ij}^{\text{Binary}} = \frac{n_{ij}}{\sqrt{n_i n_j}},$$

(2)

where $n_{ij}$ is the intersection or overlap, i.e., the count of units (e.g., songs) consumed by both users $i$ and $j$, and $n_i$ ($n_j$) is the count of units consumed by user $i$ ($j$). In this case, the cosine similarity is the fraction of the number of songs jointly consumed (i.e., the overlap) over the square root of the product of songs consumed by each user. Because the latter is equal to the geometric mean of the number of varieties consumed by the pair, the cosine similarity can be interpreted as the ratio of overlap count to average variety count. Note that this expression does not account for the positive relationship between expected similarity and variety driven by increased consumption. We turn to this issue next.

**Decomposing Similarity into Size and Content Components**

Similarity can increase because consumers expand their consumption sets, making overlap more likely, even when platforms make no recommendations and consumers have
independent tastes. To illustrate, assume a music catalog with 100 songs. When users \( i \) and \( j \) listen to songs from this catalog independently and all songs have an equal probability of 0.10 to be heard, we would expect this pair to share one song in common. Substituting this overlap into equation (2), we expect their similarity is 0.10.\(^5\) Now suppose user \( i \) adopts Spotify, and the zero marginal cost of accessing individual songs makes the consumer listen more frequently, such that her songs have a 0.20 chance to be heard. We would now expect the pair to have two songs in common, yielding an expected cosine similarity of 0.14.\(^6\)

More generally, let \( X_{ik} \) and \( X_{jk} \) for \( k = 1, \ldots, N \) be i.i.d. random variables with finite means \( E[X_{ik}] \) and \( E[X_{jk}] \) that are also independent across users, i.e., \( E[X_{ik} X_{jk}] = E[X_{ik}] E[X_{jk}] \). We show in Web Appendix B that under independence across users, the expectation of cosine similarity converges in probability to a ratio of moments:

\[
E[\text{Cosine Similarity}_{ij}] \xrightarrow{p} \frac{E[X_{ik}] E[X_{jk}]}{\sqrt{E[X_{ik}^2] E[X_{jk}^2]}}
\]

This expression holds for any distribution of \( X_{ik} \) or \( X_{jk} \). If the elements of \( X \) are binary, \( E[X_{ik}] = E[X_{ik}^2] = \hat{p}_i = n_i/N \), and \( E[X_{jk}] = E[X_{jk}^2] = \hat{p}_j = n_j/N \). Then, substituting these estimators into equation (3) gives

\[
E[\text{Cosine Similarity}_{ij}^{\text{Binary}}] \xrightarrow{p} \frac{\sqrt{n_i n_j}}{N}.
\]

Equation (4) shows that the relationship between consumption set sizes (\( n_i \) and \( n_j \)) and expected similarity is positive and nonlinear. If both users double the number of varieties consumed, their

\(^5\) The joint distribution two independent and identically distributed (binary) random variables is equal to the product of its marginal distributions; the probability of jointly consuming a song is \((0.1) \times (0.1)\); multiplication by assortment size yields the expected overlap = \((0.1) \times (0.1) \times 100 = 1\), and the Expected Cosine Similarity = \(1/\sqrt{100 \times 10} = .10\).

\(^6\) Expected overlap = \((0.2) \times (0.1) \times 100\); Expected Cosine Similarity = \(2/\sqrt{20 \times 10} = .14\)
expected similarity should double as well. If only one of the pair doubles, the expected similarity is multiplied by a factor $\sqrt{2}$, increasing by 41%, exactly as the example given above (where similarity increases by 41% from 0.10 to 0.14).

Clearly, this example shows how similarity can increase when a user’s consumption set size increases, even in the absence of any Spotify-directed search. Empirically, our strategy will be to compute these expressions for each pair of consumers and period to offset observed similarity by the similarity expected by consumption set expansion alone (equation 4).

Continuing our example above, suppose that when one member of the pair adopts Spotify, we observe pairwise similarity increases by only 30%, rather than 41% as predicted by the size effect alone. That means that the difference, 30% - 41% = -11%, is the content effect. In other words, Spotify-directed search (conditional on consumption set size) decreases similarity by 11%. This way, we decompose Spotify induced changes in listeners’ similarity into a size effect of consumption expansion, and a content effect due to coordination in music consumption. This distinction is relevant because consumer’s self-chosen expansion of listening variety is associated with positive consumer welfare effects, but similarity from redirecting listeners’ consumption away from their tastes is not.

**Accounting for Differential Popularity**

One confounding factor is the differential popularity of content. Users may flock to the same songs, increasing across-user similarity, because they have higher quality or other characteristics unrelated to any Spotify-related coordination. Furthermore, in deriving the expected similarity above, we also assumed that the probability of consumption is identical across artists, which may not be realistic.
We are not the first to note this problem. The literature on information retrieval also accounts for differential popularity when calculating textual similarity to compare documents. Popular words such as “the” and “and” are not informative because they appear in many documents (Murphy 2012). We follow the strategy common in that literature (Manning, Raghavan, and Schütze 2009), and pre-multiply our consumption data with content-specific weights. Specifically, as is customary in this literature, we weight each song by its log inverse popularity, i.e.,

$$w_k = \log \left( \frac{1}{p_k} \right) \quad \text{and} \quad p_k = \frac{\sum_i (X_{ik} > 0)}{I},$$  \hspace{1cm} (5)

where $p_k$ is the proportion of users who listen to song $k$. The more uncommon the song, the lower $p_k$ and the higher the weight, which means that the play count for this song is more informative about similarity. Conversely, the more popular the song, the lower the weight becomes, making this less important for calculating similarity. At the extreme, if everyone listens to a particular song, i.e., $p_k = 1$, the weight is zero, and its consumption does not count.

Furthermore, by down-weighting popular content and up-weighting unpopular content, the transformed data becomes more identically distributed, which is in line with our assumption of equal consumption probabilities when computing expected similarities from consumption expansion.

This particular functional log form is known as the inverse document frequency and is widely used. Elkan (2005) shows that a model where consumer choice is governed by a Dirichlet-Multinomial distribution and similarity between any two consumers is calculated using a Fisher kernel leads to an expression close to equation (5). We then adjust our measures as follows:
Weighted Cosine Similarity$_{ij} = \frac{\sum_k X_{ik} X_{jk} w_k^2}{\sqrt{\sum_k X_{ik}^2} \sqrt{\sum_k X_{jk}^2}}$ \hspace{1cm} (6)

$E[\text{Weighted Cosine Similarity}_{ij}] = \frac{\sqrt{\sum_k X_{ik}^2} \sqrt{\sum_k X_{jk}^2}}{\sum_k w_k^2}$ \hspace{1cm} (7)

These expressions nest their counterparts above when weights are equal across songs. In our empirical application, the weights are dynamic, calculated for each song and four-week time period, but our results are quantitatively similar if we use static or (one-period) lagged weights.

**Measuring Adoption**

The unit of analysis is a pair of users, $i$ and $j$, in a given four-week period $t$. We focus on the treatment effects starting at the moment of first and second adoption in the pair. In the baseline scenario, we construct step-function dummy variables which are zero until the period at which the first (and second, respectively) of the pair adopts, and are one after:

\[
\text{first\_adopt}_{ijt} = 1_{\{t \geq \tau_i \cup t \geq \tau_j\}}
\]

\[
\text{second\_adopt}_{ijt} = 1_{\{t \geq \tau_i \cap t \geq \tau_j\}}
\]

where $\tau_i$ is the adoption period of user $i$. We set $\tau_i$ to infinity, if user $i$ never adopts. If neither user adopts, these variables remain zero. If both users in the pair choose to adopt Spotify in the same period (occurring in 7% of both adopting pairs), both adoption variables are set to 1 in the same period.
DATA

Data Collection and User Sample

We use panel data collected from a third-party music recommendation service\(^7\) that monitors song-level consumption by individual users on all platforms activated by the user. The service integrates individual-level music consumption on more than 100 possible devices and clients, including offline and online, mobile, and desktop. Our sample frame consists of the service’s user base. Similar to Datta, Knox, and Bronnenberg (2018), our exact user sample was selected from this sample frame by repeatedly visiting the service's website in short intervals between April 22 and April 29, 2014. Our web scrapers collected the user-names of recently active users on the service during this period. Using the service’s application protocol interface (API), we recorded age, gender, and country for a random selection of 5,003 users,\(^8\) and retrieved their music consumption histories for 2.5 years (January 6, 2013, until August 1, 2015). From May 1, 2014, onwards, we also track users’ platform choices, by visiting users’ profile pages about every 16 minutes.\(^9\) In our data, we observe 4,033 active users, and 121 million songs played during that period.\(^10\) We aggregate the consumption data to the user-period level, where a period is four weeks.

---

\(^7\) For reasons of confidentiality, the data provider wishes to remain anonymous.
\(^8\) In the cases for which age and gender were missing (22.4% and 12.4% of all users, respectively), we first estimated age and then gender based on an auxiliary regression model for age (and a logit model for gender) using covariates capturing listening behavior (e.g., quantity, variety consumed).
\(^9\) We impose a retrieval limit of 5 pages per second (in line with the service’s API Terms of Use); that yields an approximate retrieval time for all users of 5,003 users / 5 users per second = 1,001 seconds = 16.6 minutes.
\(^10\) Because of technical reasons, we miss consumption for 27 days (out of 458 days between 1 May 2014 and 1 August 2015); missing days were not systematic. Data for 970 users (5,003-4,033) is not available, either because these users changed their privacy settings or cancelled the service during our observation period. In preparing the data, we remove all songs that were listened to less than 30 seconds (cf. Datta, Knox and Bronnenberg 2018, who have removed all songs played for less than half of their song length). For more details about the data collection, see Datta, Knox and Bronnenberg (2018).
We identify unique artists and songs using an algorithm described in Web Appendix C, so that “The Beatles” / “beatles” and “yesterday” / “Yesterday” are counted as the same artist and song, respectively. Second, we link each unique artist and song in our data set to additional databases, the open-source music encyclopedia Musicbrainz.org and the freely available Discogs database (both obtained in 2018), and the music intelligence company Echonest.com (data collected in 2016). We use these additional databases to construct our similarity measures, e.g., based on meta-characteristics such as a song’s artist. The most played artists in the raw data are Lana Del Rey, Taylor Swift, and Madonna.

Although we focus on the impact of the adoption of the leading on-demand streaming provider at the time of our study, Spotify, our data records all music consumption except FM radio. Hence, we measure consumption similarity across many platforms, not only Spotify. There are other streaming providers in the data, but their collective share is negligible.

We only observe adoption if someone who has not previously used Spotify starts using it. Because of left-truncation bias, as in Datta, Knox, and Bronnenberg (2018), we remove users who adopt Spotify less than 45 days into our observation window (including those who adopt before 1 May 2014). Also, we require potential adopters to be active before adoption on at least one of the major ownership-based platforms at the time (i.e., iTunes, Windows Media Player, Winamp, or Foobar).

This trimming of consumers in our data leaves us with 2,311 listeners, of which 561 adopt Spotify in our observation period. We observe users for up to 17 consecutive four-week periods (the average duration of observation is 13 periods). Users are, on average, 24 years old and 76% male, with 31% coming from the European Union, 20% from South America, and 10% from Canada and the US. The average number of songs heard each period is 880. Assuming the
average song length is 4 minutes, this means 59 hours per four-week period, or a little less than 2 hours per day.\textsuperscript{11} This listening is spread out across (on average) 451 unique songs from 92 unique artists.

\textit{Constructing Pairs, and Computing Similarity Metrics}

We construct pairs of users from our sample of users. There are 2,669,205 possible (undirected) pairs among 2,311 users.\textsuperscript{12} However, with some pairs, there may be no simultaneous overlap, e.g., user $i$ is active in periods 1-5, but user $j$ is only active in periods 6-17. Secondly, if there is no variation in the dependent variable (e.g., the similarity between users $i$ and $j$ is always zero), there is no within-pair change to measure, so these pairs are excluded in the estimation (Wooldridge 1999). As a result, we end up with 1,448,107 pairs (observed over 20,401,087 periods, see Table 1) for songs and 1,884,040 pairs (observed over 25,894,511 periods) for artists in the sample.

Our dependent variable is weighted cosine similarity (equation 6) measured separately for the consumption of songs and artists. Table 1 gives summary statistics for these variables as well as the average unweighted number of jointly consumed songs and artists ($n_{ij}$ in equation 2). The first row of Table 1 shows the per period average number of jointly consumed songs is 2.16, and the average number of jointly consumed artists is 1.73. The mean observed cosine similarity is 0.00262 for songs and 0.00834 for artists, and the distribution is highly skewed: 67\% of song and 53\% of artist observations are zero (i.e., have non-overlapping consumption sets).

As a comparison, Table 1 also shows expected cosine similarity (equation 7), i.e., expected similarity if consumption were independent across pairs given the total amount of

\textsuperscript{11} We use the average song length across 27 million tracks on Musicbrainz.org, winsorized at 2.5\% to account for outliers (4.00598 minutes). Consumption then is $880 \times \frac{4}{60} = 58.7$ hours per period, or $\frac{58.7}{30} = 1.96$ hours per day.

\textsuperscript{12} There are $\frac{(2311)(2310)}{2}$ possible unique pairs.
variety consumed in a given period. We see that the average expected similarity is about an order of magnitude smaller than actual similarity: 0.000644 for songs and 0.000752 for artists. The ratio of actual to expected similarity is about 8 for songs and 26 for artists. Relative to expected similarity, which assumes consumption across users is independent, actual similarity is large.

Table 2 provides model-free evidence that the average number of jointly consumed varieties, observed similarity and expected similarity vary with Spotify adoption. We compare average similarity over four samples: observations from pairs of users who never adopt (never adopters), pairs with at least one adoption, but before the event occurs (before adoption), pairs with only one adopter and pairs with two adopters, both after adoption. The average number of jointly consumed songs and average similarity is higher for adopters before they adopt than never adopters, which suggests some selection in who joins Spotify and reinforces our decision to measure pre-existing similarity as a baseline.

We focus on adopters. When one user adopts, jointly consumed songs decreases from 2.70 to 2.37. However, when both users adopt Spotify, the average number of co-consumed songs increases to 3.58. We see the same pattern in terms of observed cosine similarity: in terms of songs, it decreases by 10% (from 0.00318 to 0.00285) when one member joins but increases by 45% (from 0.00285 to 0.00414) when both join. In other words, after both members of the pair adopt, their listening is 30% higher than before either of them adopted (from 0.00318 to 0.00414). We see the same result using artists. The model-free evidence shows that similarity is higher when both members have adopted.
But to what extent is this increase due to adopters increasing their consumption set as opposed to adopters switching to content that is more similar to each other? Table 2 also reports expected cosine similarity given the size of the consumption set and the ratio of actual to expected cosine similarity. If similarity increases were due to variety expansion, we would expect the ratio to remain constant. Alternatively, if similarity increases were the result of content switching, we would expect to see the ratio increase.

Table 2 shows that it goes from 10 before adoption, to 8 when one member adopts, back to the pre-adoption level of 10 when both members have adopted. Similarity, the ratio for artists is about 35 before adoption, drops to 28 when one member adopts, but returns to 35 when both have adopted. So, the view that similarity increases are due to variety expansion—the ratio of observed to expected is about the same for pairs before adoption and when both have adopted—is supported by the model-free evidence. We investigate this more systematically with the model, controlling for pre-existing similarity and time trends with fixed effects in the next section.

**MODEL**

The modal observed weighted cosine similarity is 0, as most pair-period observations have non-overlapping consumption sets. Instead of adding an arbitrary (positive) number to the dependent variable before taking its logarithm, we opt for a Poisson fixed effects estimator (also called the Poisson pseudo-maximum likelihood model in the economic literature, see Santos and Tenreyro 2006 and 2011), which has an appealing log-linear interpretation, yielding estimates that are a percentage change from pre-Spotify similarity, yet where zero is a valid outcome. Wooldridge (1999) shows that this model has very appealing robustness properties: it is valid for any distribution of the dependent variable (not just count data) provided it is non-negative; it can be over- or under-dispersed, and is stable when the proportion of zeros is large (Santos and
Tenreyro 2011). He also shows that under very general conditions, this model yields consistent estimates, e.g., even with arbitrary dependence over time.

We model the weighted cosine similarity (equation 6) between user \(i\) and \(j\) at period \(t\) as a function of pair-specific fixed effects \((\delta_{ij})\) capturing baseline similarity between users, a set of fixed effects \((\gamma_t)\) capturing common time shocks to similarity (e.g., the Holiday Season), two treatment effects for when the first user \((\beta_1)\) or second user in the pair \((\beta_2)\) adopts a streaming service (equation 8), and, an offset (equation 7) that controls for variety expansion, i.e., the expected pairwise similarity under independence \((O_{ijt})\):

\[
\text{Similarity}_{ijt} = \exp(\delta_{ij} + \gamma_t + \beta_1 \text{first}_{\text{adopt}}_{ijt} + \beta_2 \text{second}_{\text{adopt}}_{ijt} + \ln(O_{ijt})) + \epsilon_{ijt} \tag{9}
\]

We note that the first-order conditions for the treatment effects are the same as the fixed effects Poisson estimator. We estimate two versions of equation (9): one with the offset controlling for variety expansion to estimate the identity component conditional on consumption set size, and one without the offset, to estimate the total effect of adoption on similarity.

Controlling for baseline pairwise similarity is important to separately identify the effects of adoption. For example, users may be friends with each other or share similar friends; they may also share similar backgrounds or demographics. We remove this baseline similarity with the pair-specific intercept \(\delta_{ij}\).

We also control for dependence across pairs over time in our standard errors, as our data are constructed using all possible pairs of users, wherein two pairs can share the same user. Such observations are not independent because a shock to the common user’s tastes will affect all pairs of which that user is a member. Ignoring these dependencies would bias our standard errors downwards. We correct for this cross-sectional pairwise dependence by using multiway clustered standard errors (Cameron, Gelbach, and Miller 2011), effectively clustering on both members of
the pair. We implement the multiway-clustered standard errors using the method proposed by Correia, Guimaraes, and Zulkin (2019).

RESULTS

Spotify Adoption Increases Overall Consumption Similarity

Does Spotify adoption increase or decrease overall consumption similarity? Table 3, columns 1 and 3, present estimates from equation (9) without the offset, using cosine similarity, computed using weighted binary listening measures, with the weights based on period-specific popularity, for songs and artists. We preface the discussion by noting that our results are robust to measuring similarity along an intensive margin, i.e., the log of play counts instead of binary measures, using the Jaccard set similarity instead of cosine similarity, and to using static popularity weights or lagged weights (see Web Appendix D for the results).

[Insert Table 3 about here.]

Column 1 shows the results for songs. When one user adopts Spotify (“first adoption”), pairwise similarity to the other user who has not adopted increases permanently by 9% (= exp(.090) – 1). If the other member of the pair also adopts (“second adoption”), similarity increases a further 14% (= exp(0.130) – 1). These estimates imply that total similarity increases by 25% (= exp(.090 +.130) – 1) when both members in the pair join Spotify.

Column 3 shows the equivalent effects if we measure similarity in terms of artists rather than songs. When one user adopts Spotify, similarity to the non-adopted user increases by 11% (= exp(.010) – 1); if both users adopt, similarity increases another 15% (= exp(.014) – 1), implying total similarity increases by 27% (= exp(.010 +.014) – 1) relative to pre-adoption levels. Hence, these results suggest that Spotify is substantially increasing overall similarity in
listening across consumers, consistent with the model-free evidence in Table 2 and observations in the popular press (e.g., Guardian 2019).

**Conditional on Consumption Set Expansion, Spotify Adoption Decreases Similarity**

How much of this increase in consumption similarity is due to adopters switching to more similar varieties rather than simply choosing more variety overall? After accounting for consumption set size expansion, does Spotify make users’ consumption more or less individualized? We now focus on the content component by including the offset in equation (9). Interestingly, when we control for size expansion this way, the direction of the effect becomes *negative*, implying that users consume *less*, not more, similar content after controlling for consumption set size.

Column 2 shows the results for songs. When the first user adopts Spotify, pairwise song similarity decreases by 4% (=exp(-.036) – 1). Another way of interpreting this is that similarity is 4% lower than we would expect from consumption set expansion alone. The results suggest that similarity decreases further if the second user adopts, but the effect is smaller and insignificant.

For the similarity in the consumption of artists (as opposed to songs), we see even stronger results. Column 4 of the table reports that similarity decreases by 9% (=exp(-.09) – 1) if one user adopts and that it drops another 9% (=exp(-.09) – 1) if both users adopt. These numbers, therefore, imply that after controlling for expanding variety, two adopters are 16% (=exp(-.09 -.09) – 1) less similar to each other after adopting Spotify than before it. Hence, the content component goes in a *different* direction and is smaller than the total effect. After accounting for consumption set expansion, users’ consumption becomes *more individualized*.

This also means that the effect from only consumption expansion on similarity is larger than the combined effect of both the size and content components. While not shown in the tables,
that effect is straightforward to calculate. For example, for artists, the size expansion effect for the first adopter is 0.19 (= 0.10 – (-0.09)) and for the second adopter 0.23 (= 0.14 – (-0.09)), which translates into a 21% (=exp(.19)-1) and 26% (=exp(.23)-1) increase in similarity, respectively. In sum, the consumption set expansion effect is positive and larger in magnitude than the total effect. This indicates that the overall increase in similarity is entirely driven by consumption set expansion, not reallocating choices.

**The Effects are Stronger for Heavier (vs. Lighter) Users**

We now focus on the content component of consumption similarity: where are these negative effects on consumption similarity strongest? From our introduction, we expect that Spotify infers users’ tastes from consumption histories. This suggests that the intensity of Spotify usage moderates these effects. In particular, we would expect that heavier Spotify users experience larger negative effects than light Spotify users because heavier users receive more personalized playlists by virtue of their heavier usage. We test this hypothesis by classifying users based on their consumption intensities. We do so by calculating for each user their average per-period Spotify play count (usage) and comparing this to the median per-period usage. We then estimate the regression specification in equation (9) separately for heavy and light users.

[Insert Figure 1 about here.]

Figure 1 shows the Spotify-adoptions effects in terms of the percentage change in weighted cosine similarity. If the first adopter is a heavy user, the content component of the adoption effect, i.e., song similarity controlling for consumption expansion, falls by 6%; if he or she is a light user, there is approximately no change (the difference in first adoption effects between heavy and light users is significant at $p = 0.014$). There are no differences in second adopter effects for songs (but there were no significant effects overall for songs).
For artists, we see clearly that adoption effects are more negative for heavy than light users for both first ($p < .001$) and second adopters ($p = 0.013$). In summary, after accounting for consumption expansion, Spotify adoption causes users to become less similar, with pairs of heavier users becoming more dissimilar than pairs of light users.

**The Effects Become Stronger Over Time Since Adoption**

To study the evolution of similarity (conditional on size) over time, we estimate equation (9), but we replace our fixed first and second adopter treatment effects with adoption effects that are specific to time since adoption: 0 periods after (either first or second) adoption, e.g., the period of adoption, 1-2 periods after adoption, 3-5, 6-8, 9-11, and 12+ periods after adoption (we observe at maximum 17 periods, and a period is 4 weeks). These six categories were chosen because they lead to an approximately equal number of observations per group. Figure 2 shows these adoption coefficients as a function of time since adoption, for similarity in consumption of songs (left) and artists (right).

[Insert Figure 2 about here]

Overall, pairwise consumption similarity conditional on size (i.e., the content component) decreases over time since Spotify adoption. However, the standard errors of the effects are large relative to changes over time. The decrease over time is more pronounced when it concerns second adopter effects—if both members join Spotify—than first adopter effects. If one member joins Spotify, song similarity immediately drops 3%; 12+ periods later similarity is down 4%, one percentage point less, but this difference is not significant. If both members join Spotify, song similarity drops 1% immediately and is down 7% after 12 periods, a marginally significant difference ($p = 0.11$). In the case of artist similarity, if one member adopts similarity drops 9% immediately and is down 11% a year later, but this difference is not significant. If both members
adopt, similarity drops immediately by 7%, but is down 15%, more than twice the initial amount, a year later. This difference is significant ($p = 0.0079$).

Hence, when both users are Spotify adopters, we find evidence—more pronounced for artists than songs—that Spotify makes users increasingly dissimilar over time after controlling for size. Like our results on usage intensity causing dissimilarity, this result is consistent with the idea that platforms require individual listening histories to make personalized recommendations: the more time users spend on the platform, the more customized, and hence the more different their consumption is.

**CONCLUSION AND IMPLICATIONS**

Given massive assortments—in the case of Spotify over six lifetime’s worth of listening—it would seem that platforms have enormous scope to influence users’ consumption decisions. Appearing on a major same-for-everyone playlist like Today’s Top Hits increases consumption by nearly 20 million streams and is worth between $116,000 and $163,000 (Aguiar and Waldfogel 2018a). And recommender systems that are behind customized playlists have well-known popularity biases (Fleder and Hosanagar 2009). Hence a significant concern of critics and policy-makers is taste homogenization due to, in particular, the power of platforms to push content against the tastes of consumers (e.g., Guardian 2019; Economist 2018).

We construct a pairwise panel of listening and investigate similarity around the time when one and both members of the pair join a streaming platform. We find that adopting Spotify increases similarity in consumption across consumers, in line with the observation that Spotify is homogenizing tastes. However, we view the same overall result very differently. Our results indicate this is due to expanding yet individualizing variety. Most consumers like variety in music. Because variety is free at the margin on Spotify, users consume more of it (Datta, Knox,
and Bronnenberg 2018), which raises listening similarity independent of any platform-induced coordination. This increase is associated with something most consumers would find valuable. After accounting for the effects of consumption expansion, we find that Spotify makes consumers develop more individual listening histories, and this effect becomes more pronounced with usage intensity and length of usage experience. Although our data do not allow us to make statements about welfare, this effect, too, is likely something consumers find valuable. Thus, we conclude there is little evidence to suggest that Spotify makes consumers more similar in a (for consumers) disadvantageous way. We also find that the overall level of similarity is low. Hence, one implication of our study is the proposition that platforms are harming consumers by making them more similar is misleading.

The result that – conditional on consumption set size – users become less similar is also different from Hosanagar et al. (2014), who find positive similarity effects in their “product-mix” component. This finding is likely due to the business models of streaming platforms in which variety is free at the margin. In ownership platforms like iTunes, the fact that variety is costly may skew users’ choices to more popular content in the presence of recommender systems.

Consumption set expansion as the primary driver of observed increasing similarity is also good news for musicians, producers, and advertisers. For musicians and producers, users listening to more variety means a higher chance of appearing in users’ expanded consumption sets. We see these effects after controlling for heterogeneous popularity, so it is not the case that these effects are driven by increased consumption of superstars (i.e., artists) or hits (i.e., individual songs), in contrast to the view that platforms lead to winner-take-all markets.

For advertisers, increased similarity—holding variety constant—means fewer opportunities to target based on consumption. In the extreme case, if platform adoption ensured
that all users listen to the same bundle of variety, using consumption data, marketers would have only minimal targeting opportunities, i.e., everyone or no one. Holding size constant, we find Spotify makes consumption less, not more, similar, which means more opportunities for segmenting the market. Furthermore, we find that variety is not constant, but increasing, which also increases the value to advertisers. One caveat is that we cannot make statements about whether this is worse for marketers: that would require modeling consumer tastes, which are what drives the value of the data to marketers.

There are several caveats to our study. First, because we study music distribution, we cannot say much about other industries. But given the importance of globally curated lists (e.g., New York Times Best Sellers) and personalized recommendations (e.g., Amazon.com’s “Top Selected Products and Reviews”), we expect our results are relevant in other markets such as books and TV shows where consumers like variety.
REFERENCES


Table 1: Summary statistics

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>10th percentile</th>
<th>90th percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Song-level</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jointly consumed songs</td>
<td>20,401,087</td>
<td>2.16</td>
<td>6.94</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>Observed cosine similarity</td>
<td>20,401,087</td>
<td>.00262</td>
<td>.00816</td>
<td>0</td>
<td>.00706</td>
</tr>
<tr>
<td>Expected cosine similarity</td>
<td>20,401,087</td>
<td>.000644</td>
<td>.00548</td>
<td>.000144</td>
<td>.001310</td>
</tr>
<tr>
<td>Ratio observed-to-expected</td>
<td>20,401,087</td>
<td>8.11</td>
<td>96.9</td>
<td>0</td>
<td>13.6</td>
</tr>
<tr>
<td><strong>Artist-level</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jointly consumed artists</td>
<td>25,894,511</td>
<td>1.73</td>
<td>3.95</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Observed cosine similarity</td>
<td>25,894,511</td>
<td>.00834</td>
<td>.01800</td>
<td>0</td>
<td>.024200</td>
</tr>
<tr>
<td>Expected cosine similarity</td>
<td>25,894,511</td>
<td>.000752</td>
<td>.000765</td>
<td>.000147</td>
<td>.001580</td>
</tr>
<tr>
<td>Ratio observed-to-expected</td>
<td>25,894,511</td>
<td>26.4</td>
<td>319</td>
<td>0</td>
<td>49.9</td>
</tr>
</tbody>
</table>

Notes: Summary statistics for jointly consumed units, observed similarities, expected similarities (under independence, i.e., the size component of similarity), and the ratio of observed similarities to expected similarities are calculated on a sample of 2,509,649 user-pairs, created from all possible pairs of 2,247 (561 adopters and 1,686 nonadopters), observed over 17 4-week periods starting May 29, 2014, yielding a total of 32,572,576 observations. Actual pairs and observations differ depending on the type of consumption measured. Pairs with a single observation or no within-pair variation in similarity are dropped; the unit of analysis is the pair-period (4 weeks). Weights used for all similarity measures, which are inversely proportional to song and artist popularity.
### Table 2: Average Similarity by Adoption: Within-Adopter Means

<table>
<thead>
<tr>
<th></th>
<th>Never adopters</th>
<th>Adopters</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Before adoption</td>
</tr>
<tr>
<td>Song-level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jointly consumed songs</td>
<td>1.71</td>
<td>2.7</td>
</tr>
<tr>
<td>Observed cosine similarity</td>
<td>.002150</td>
<td>.003180</td>
</tr>
<tr>
<td>Expected cosine similarity</td>
<td>.000651</td>
<td>.000608</td>
</tr>
<tr>
<td>Ratio observed-to-expected</td>
<td>6.92</td>
<td>10.4</td>
</tr>
<tr>
<td>Artist-level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jointly consumed artists</td>
<td>1.45</td>
<td>1.96</td>
</tr>
<tr>
<td>Observed cosine similarity</td>
<td>.006990</td>
<td>.009800</td>
</tr>
<tr>
<td>Expected cosine similarity</td>
<td>.000774</td>
<td>.000681</td>
</tr>
<tr>
<td>Ratio observed-to-expected</td>
<td>22.1</td>
<td>34.7</td>
</tr>
</tbody>
</table>

Notes: Estimates are calculated on a sample of 2,509,649 user-pairs, created from all possible pairs of 2,247 (561 adopters and 1,686 nonadopters), observed over 17 4-week periods starting May 29, 2014, yielding a total of 32,572,576 observations. Actual pairs and observations differ depending on the type of consumption measured. Pairs with a single observation or no within-pair variation in similarity are dropped; the unit of analysis is the pair-period (4 weeks). We calculate the average similarity for four groups of pairs: those who never adopt, pairs where at least one adopts, before adoption, pairs where one adopts on or after adoption, and pairs with both adopters after adoption. Weights used for all similarity measures, which are inversely proportional to song and artist popularity.
Table 3: Adoption effects on consumption similarity

<table>
<thead>
<tr>
<th></th>
<th>(1) Total consumption similarity</th>
<th>(2) Content component of consumption similarity</th>
<th>(3) Total consumption similarity</th>
<th>(4) Content component of consumption similarity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Song-level</td>
<td>Artist-level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>First adopt.</td>
<td>0.090***</td>
<td>-0.036**</td>
<td>0.100***</td>
<td>-0.090***</td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.013)</td>
<td>(0.011)</td>
<td>(0.012)</td>
</tr>
<tr>
<td>Second adopt.</td>
<td>0.130***</td>
<td>-0.022</td>
<td>0.140***</td>
<td>-0.090***</td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.020)</td>
<td>(0.016)</td>
<td>(0.018)</td>
</tr>
<tr>
<td>LL</td>
<td>-286,427</td>
<td>-285,564</td>
<td>-961,485</td>
<td>-962,807</td>
</tr>
<tr>
<td>Pairs of users</td>
<td>1,448,107</td>
<td>1,448,107</td>
<td>1,884,040</td>
<td>1,884,040</td>
</tr>
<tr>
<td>Observations</td>
<td>20,401,087</td>
<td>20,401,087</td>
<td>25,894,511</td>
<td>25,894,511</td>
</tr>
</tbody>
</table>

Notes: Regression with multiway-clustered standard errors by each member of user-pair in parentheses. Estimates are calculated on a sample of 2,655,187 user-pairs, created from all possible pairs of 2,311 (561 adopters and 1,750 nonadopters), observed over 17 periods starting May 29, 2014, yielding a total of 32,718,114 user-pairs. Actual pairs and observations differ depending on the type of consumption measured (i.e., song versus artist); user-pair- and period-specific fixed effects are used, and the unit of analysis is the pair-period. The dependent variables are (weighted) cosine similarities of binary listening indicators, equal to 1 if a consumer listens to a particular song (columns 1-2) or artist (columns 3-4). The independent variables are adoption dummy variables that are 1 if the period is on or after the first (second) user in the pair adopts Spotify, and zero otherwise. Total consumption similarities (columns 1 and 3) are based on equation (9); consumption similarity from the content component (columns 2 and 4) are based on equation (10). The coefficient of the offset $O_{ijt}$ is logically equal to one and not shown in the table.

* $p < 0.10$, † $p < 0.05$, ‡ $p < .01$, §§ $p < .001$. 
Figure 1: Content Component of Consumption Similarity is Stronger for Heavy Users

Notes: Figure show estimates for first and second adoption indicator variables from equation (10), estimated separately for heavy versus light users of Spotify (based on a median split on monthly Spotify usage per user) for consumption similarity in the number of songs and artists listened.
Figure 2. Consumption Similarity from Discovery over Time

Notes: We replace the first and second adoption indicator variables in equation (10) with indicator variables, capturing the time since adoption of the first and second adopter (measured in 4-week periods), for similarity in the number of songs and artists listened.
WEB APPENDIX A: SIMULATIONS OF SIMILARITY AND RESULTING INEQUALITY

We simulate binary consumption decisions of 250 users over 10000 artists under differing levels of cosine similarity for all user-pairs (here denoted $\rho$). Each user has a .001 i.i.d. Bernoulli probability of consuming each artist (so on average, each user consumes 100 artists out of the total 10000. We consider three scenarios: independence ($\rho=0$), positive ($\rho=0.0001$), more positive ($\rho=0.0002$). We display the results below, where each dot is a play. Concentrating listening corresponds to vertical “streaks” in the figure below. While there are hardly any streaks in the leftmost graph, under independence, there are several in the right two graphs, as similarity is increasing. We also show the Herfindahl index across artists is increasing as we move from left to right.

**Figure A1.** *Similarity is positively related to concentration*


WEB APPENDIX B: GENERAL EXPRESSION FOR EXPECTED COSINE SIMILARITY UNDER INDEPENDENCE

Let $X_{ik}$ and $X_{jk}$ be i.i.d. random sequences over $k$ with finite expectations, i.e., $E[X_{ik}]^{1+\delta} < \infty$ and $E[X_{jk}]^{1+\delta} < \infty$ for some $\delta > 0$. We assume that these random variables are independent of each other.

$$E[X_{ik}X_{jk}] = E[X_{ik}] E[X_{jk}]$$  \hspace{1cm} (A1)

If we further assume that $X_{ik}$ and $X_{jl}$ are independent of each other for $k \neq l$, then the product of two elements, $X_{ik} X_{jk}$, is also an i.i.d. sequence.

Slutsky’s theorem says that the probability limit of a ratio is the ratio of the probability limits:

$$\text{plim} \frac{X_i X'_j}{\|X_i\| \|X_j\|} = \frac{\text{plim} X_i X'_j}{\text{plim} \|X_i\| \text{plim} \|X_j\|}$$  \hspace{1cm} (A2)

For the numerator, the weak law of large numbers (WLLN) says that the mean product converges in probability to the expected value of the product:

$$\frac{1}{N} \sum_{k=1}^{N} X_{ik} X_{jk} \overset{p}{\rightarrow} E[X_{ik}X_{jk}]$$  \hspace{1cm} (A3)

Under independence, the expectation of the product is the product of the expectation (A1).

For the first term in the denominator, because $X_{ik}$ is i.i.d., so is any function of it, including the L2 norm. If $E[X_{ik}]^{2+\delta} < \infty$ for some $\delta > 0$, by the WLLN and the continuous mapping theorem:

$$\sqrt{\frac{1}{N} \sum_{k=1}^{N} X_{ik}^2} \overset{p}{\rightarrow} \sqrt{E[X_{ik}^2]}$$  \hspace{1cm} (A4)

The same holds for the second term in the denominator. Then the cosine similarity converges in probability to the expression in the main text:

$$\text{plim} \frac{X_i X'_j}{\|X_i\| \|X_j\|} = \frac{E[X_{ik}] E[X_{jk}]}{\sqrt{E[X_{ik}^2] E[X_{jk}^2]}}$$  \hspace{1cm} (A5)

This is the general formula, which does not rely on any particular distribution. We also test these convergence results using simulated data, for which we need to assume a specific
distribution. In Figure A1 we show the result of such a simulation, where $X_{ik}$ and $X_{jk}$ are binary with $P(X_{ik} = 1) = p = 0.05$ and $P(X_{jk} = 1) = q = 0.1$ and $N = 10000$, and the total number of simulations is 1000. The figure compares the histogram of simulated similarities with the theoretically derived one in red. The empirical average is the same to two decimal points.

**Figure B1. Small sample convergence of cosine similarity to probability limit**
WEB APPENDIX C: ALGORITHM TO IDENTIFY UNIQUE ARTISTS AND SONGS

The aim of this algorithm is twofold. First, we identify unique artists and songs, correcting labels so that “The Beatles”/“beatles” and “yesterday”/“Yesterday” are counted as one artist and song, respectively. Second, we link each unique artist and song in our data set to additional databases—the open-source music encyclopedia Musicbrainz.org (replication of SQL database, data from July 2018), the freely available Discogs database (raw XML data files dated 1 July 2018), and the music intelligence company Echonest.com (data obtained in 2016), which powers Spotify’s music catalog—to construct our variety measures (e.g., based on meta-characteristics such as an artist’s genre, or a song’s release date).

Identifying Unique Artists

Step 1) Find and validate MBIDs (Musicbrainz IDs) for each artist in the raw service data

The raw service data (obtained in 2014-2015) contains information on each artist’s name as used on the service platform, and—if available—an artist’s Musicbrainz ID (MBID). Some MBIDs have been deprecated as they erroneously referred to artists with already existing identifiers; we match these cases to their updated MBIDs via Musicbrainz’ redirection table (table artist_gid_redirect). There are many artists without MBIDs in the raw service data; we look up missing MBIDs by matching artist names as obtained on the service platform to cleaned artist names from Musicbrainz’s main artist table (table artist), and a table recording alternative names for artists (artist_alias; e.g., Musicbrainz lists alternative spellings for artist “deadmau5” as “Dead Mouse”, “Deadmau”, “Deadmaus”, “Deadmouse”, “Deadmru5”, “Deadmau 5”, etc.).

Step 2) Find and validate Discogs Artist IDs for each artist in the raw service data

Next, we establish a linkage to the Discogs database. We first verify whether a valid Discogs ID is available with Musicbrainz; we match the remaining artists using cleaned artist names14 as obtained from the raw artist XML files at Discogs.

Step 3) Find and validate EchoNest IDs for each artist in the raw service data

Next, we use a combination of MBIDs and clear-text names to establish a linkage with Echonest. We perform our matching procedure as follows:

1) For each artist in the service data set, obtain the corresponding Echonest ID by querying Echonest for an artist’s MBID or updated MBID.

---

13 We match on strictly lower-case names with removed leading and trailing spaces (e.g., to match “Beatles” and “beatles”). For artists and songs that cannot be matched via their names directly, we first perform matches retaining only alphanumeric characters (a-z, A-Z, 0-9, and spaces; e.g., to combine “Chef special” with “Chef special”), and then correct for common spelling variants (e.g., & versus “and” in “Mumford & Sons” and “Mumford and Songs”, missing articles in “Flying Burrito Brothers” and “The Flying Burrito Brothers”). Artists and songs with names starting with the term “various” are excluded from text matching (as these typically refer an unspecified collection of artists/songs).

14 See footnote 13.
2) For each unmatched artist, perform a fuzzy match to obtain an artist’s Echonest ID, using Echonest’s Search API based on an artist’s clear-text name in the raw service data.
   a. If an artist’s MBID has been identified earlier, all search results returned by EchoNest holding conflicting MBIDs are removed from the set of potential matches (e.g., if a user listened to comedian Bob Marley, then the Jamaican reggae artist holding the same name would not be considered a viable match). For an artist without any MBID, the set of potential matches is left unchanged.
   b. Consider those records matched for which the restricted Damerau-Levenshtein distance between the artist name as recorded at the service, and the artist name as recorded by EchoNest is less or equal than 5. In case there are multiple candidate matches in the EchoNest search results, match to the artist with the lowest string distance (e.g., “Bob Marley & The Wailers” is matched to “Bob Marley & The Wailers” instead of “Bob Marley”).
   c. If an artist’s MBID is missing at the service, but available for the matching record at EchoNest, store the MBID as recorded at EchoNest in the matched artist data.

At the end of step 3, we are left with artists that can be identified by either their MBID, EchoNest ID, Discogs ID, or any combination of those.

**Step 4) Find unique artists in unmatched data to account for artist collaborations**

The set of artists without any match to either Musicbrainz, Discogs, or EchoNest are mostly collaborating artists. We account for collaborations among unmatched artists (e.g., “Patrice, Keziah Jones”) by retaining only the foremost artist (“Patrice”) of an artist’s cleaned clear-text name, and fill in corresponding identifiers from previously matched data, if available. Remaining artists for which no identifier at Musicbrainz, Discogs, or EchoNest is available are collapsed by their cleaned clear-text name.

After processing the raw service data, we are left with 806,006 unique artists, out of which 630,682 unique artists (78.2%) have a link to at least one meta-database (Musicbrainz, Discogs, or EchoNest). 559,210 unique artists have matches to either Musicbrainz or Discogs – two databases with data available at the song level which use in the next step. Matched artists to any of the three meta databases account for 99.0% of all song plays in our data. 176,324 artists are not matched to any additional database.

**Identifying Unique Songs**

The raw service data contains information on each songs’ name, but no universal identifier that can be readily used to cross-reference the data with Musicbrainz.org or Discogs (the data by EchoNest is only available at the artist level, and not at the song level). Therefore, we identify unique songs by matching their clear-text names within each artist already matched to Musicbrainz.org or Discogs, to song-level information from Musicbrainz and Discogs. Specifically, we use as input data to our matching procedure a list of song names and associated

---

15 We implement this by removing text after any of the following collaboration qualifiers: feat, ft, featuring, vs, versus, with, dash (-), slash (/), semicolon (;), plus (+), and (&, and), comma (,). Cleaning of artist names is performed as explained in footnote 13.

16 See footnote 13.
artist IDs from Musicbrainz (table `tracks`), and Discogs (parsed XML file for releases). Remaining songs for which no identifier at Musicbrainz or Discogs is available are collapsed by their cleaned clear-text name.\textsuperscript{17}

The database contains 9.3m unique songs; 5.3m songs have corresponding artist linkages to Musicbrainz or Discogs (collectively accounting for 84\% of all song plays in the data).

\textsuperscript{17} Unlike earlier, we do not remove text in brackets or after a dash ("–") to account for a song’s remixes (e.g., Curtis Mayfield’s “Move on up – Extended Version” or Charley Winston’s “Too Long (Radio Edit)”).
**WEB APPENDIX D: ROBUSTNESS CHECKS**

Table D1. Adoption effects on similarity  
(for different operationalizations of similarity)

<table>
<thead>
<tr>
<th></th>
<th>Total consumption similarity</th>
<th>Content component of consumption similarity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>First adopt.</td>
<td>Second adopt.</td>
</tr>
<tr>
<td><strong>Song-level</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cosine similarity (extensive margin)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dynamic weights (focal model)</td>
<td>0.090***</td>
<td>0.135***</td>
</tr>
<tr>
<td>Lagged weights</td>
<td>0.086***</td>
<td>0.124***</td>
</tr>
<tr>
<td>Static weights</td>
<td>0.078***</td>
<td>0.119***</td>
</tr>
<tr>
<td>Cosine similarity (intensive margin)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dynamic weights</td>
<td>0.096***</td>
<td>0.130***</td>
</tr>
<tr>
<td>Lagged weights</td>
<td>0.092***</td>
<td>0.118***</td>
</tr>
<tr>
<td>Static weights</td>
<td>0.085***</td>
<td>0.114***</td>
</tr>
<tr>
<td>Jaccard similarity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dynamic weights</td>
<td>0.097***</td>
<td>0.145***</td>
</tr>
<tr>
<td>Lagged weights</td>
<td>0.094***</td>
<td>0.136***</td>
</tr>
<tr>
<td>Static weights</td>
<td>0.090***</td>
<td>0.135***</td>
</tr>
</tbody>
</table>

| **Artist-level**         |              |               |              |               |
| Cosine similarity (extensive margin) |              |               |              |               |
| Dynamic weights (focal model) | 0.100*** | 0.135*** | -0.090*** | -0.090*** |
| Lagged weights           | 0.095*** | 0.125*** | -0.097*** | -0.091*** |
| Static weights           | 0.095*** | 0.131*** | -0.101*** | -0.105*** |
| Cosine similarity (intensive margin) |              |               |              |               |
| Dynamic weights          | 0.093*** | 0.117*** | -0.070*** | -0.081*** |
| Lagged weights           | 0.088*** | 0.105*** | -0.075*** | -0.084*** |
| Static weights           | 0.089*** | 0.109*** | -0.077*** | -0.094*** |
| Jaccard similarity       |              |               |              |               |
| Dynamic weights          | 0.103*** | 0.139*** | -0.064*** | -0.065*** |
| Lagged weights           | 0.099*** | 0.129*** | -0.065*** | -0.064*** |
| Static weights           | 0.102*** | 0.136*** | -0.064*** | -0.069*** |

* p < 0.10,  * p < 0.05,  ** p < .01,  *** p < .001.
Table D2. Adoption effects for heavy versus light users on the content component of similarity (for different operationalizations of similarity)

<table>
<thead>
<tr>
<th></th>
<th>Below median Spotify usage</th>
<th>Above median Spotify usage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>First adopt.</td>
<td>Second adopt.</td>
</tr>
<tr>
<td>Song-level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cosine similarity (extensive margin)</td>
<td>-0.011</td>
<td>-0.027</td>
</tr>
<tr>
<td>Cosine similarity (intensive margin)</td>
<td>-0.005</td>
<td>-0.029</td>
</tr>
<tr>
<td>Jaccard similarity</td>
<td>-0.002</td>
<td>-0.012</td>
</tr>
<tr>
<td>Artist-level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cosine similarity (extensive margin)</td>
<td>-0.041**</td>
<td>-0.051*</td>
</tr>
<tr>
<td>Cosine similarity (intensive margin)</td>
<td>-0.030*</td>
<td>-0.045*</td>
</tr>
<tr>
<td>Jaccard similarity</td>
<td>-0.029**</td>
<td>-0.038*</td>
</tr>
</tbody>
</table>

*p < 0.10, *p < 0.05, **p < .01, ***p < .001.

Note: Default (weighted) metrics used throughout.
Figure D1. Content Component of Similarity over Time

(for different operationalizations of similarity)