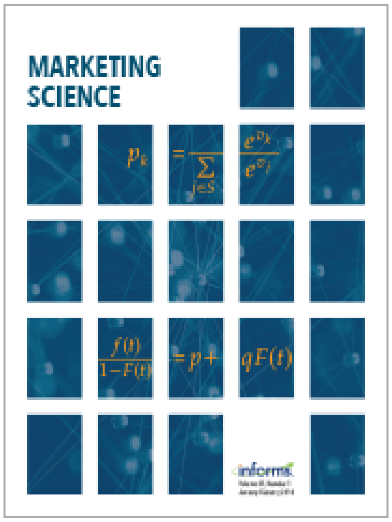


Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Changing Their Tune: How Consumers′ Adoption of Online Streaming Affects Music Consumption and Discovery

Hannes Datta, George Knox, Bart J. Bronnenberg

http://pubsonline.informs.org/journal/mksc/

MARKETING SCIENCE
Vol. 37, No. 1, January–February 2018, pp. 5–21
ISSN 0732-2399 (print), ISSN 1526-548X (online)

Changing Their Tune: How Consumers' Adoption of Online Streaming Affects Music Consumption and Discovery

Hannes Datta,[a] George Knox,[a] Bart J. Bronnenberg[a,b]

[a] Tilburg University, 5000 LE Tilburg, Netherlands; [b] Centre for Economic Policy Research, London EC1V 0DX, United Kingdom

Contact: h.datta@tilburguniversity.edu, http://orcid.org/0000-0002-8723-6002 (HD); g.knox@tilburguniversity.edu (GK); bart.bronnenberg@tilburguniversity.edu (BJB)

Received: February 9, 2016
Revised: October 19, 2016; February 13, 2017
Accepted: March 15, 2017
Published Online in Articles in Advance: September 11, 2017

https://doi.org/10.1287/mksc.2017.1051

Abstract. Instead of purchasing individual content, streaming adopters rent access to libraries from which they can consume content at no additional cost. In this paper, we study how the adoption of music streaming affects listening behavior. Using a unique panel data set of individual consumers' listening histories across many digital music platforms, adoption of streaming leads to very large increases in the quantity and diversity of consumption in the first months after adoption. Although the effects attenuate over time, even after half a year, adopters play substantially more, and more diverse, music. Relative to music ownership, where experimentation is expensive, adoption of streaming increases new music discovery. While repeat listening to new music decreases, users' best discoveries have higher play rates. We discuss the implications for consumers and producers of music.

History: Avi Goldfarb served as the senior editor and Catherine Tucker served as associate editor for this article.

Funding: The authors acknowledge financial support from the Netherlands Foundation for Science [NWO 453-09-004] and from the Marketing Science Institute [MSI 4-1854].

Supplemental Material: Data and the online appendix are available at https://doi.org/10.1287/mksc.2017.1051.

Keywords: digital distribution • online streaming • entertainment industry • music consumption • variety

1. Introduction

Copyright-related industries have suffered as new digital technologies have disrupted their revenue models. One such disruptive technology that is currently taking over the music industry is streaming. Streaming allows consumers unlimited access to a vast library of content at a fixed monthly payment. In 2015, it became the single largest source of music industry revenues in the United States (Friedlander 2016). A similar shift from ownership-based to streaming-based business models is taking place in other copyright-related industries (e.g., in particular, movies, games, and books).

As with research on file sharing, the rise of streaming has triggered a discussion among researchers about its impact on aggregate demand and producers' revenues. Using song-level digital sales, Aguiar and Waldfogel (2015) find that streaming displaces ownership-based downloads. In a survey panel, Wlömert and Papies (2016) show that free, ad-supported streaming services cannibalize demand from other channels, but since revenues from paid subscriptions more than offset this effect, streaming positively affects sales. Aguiar (2017) documents that ad-supported streaming of music increases visits to legal and illegal downloading websites among heavy users.

Missing from the literature is an account of how streaming affects consumption at the individual level. We contribute to this literature in three ways.[1] We first ask to what extent streaming generates additional music consumption rather than displacing consumption from other platforms. Access to a wide variety of content on a streaming platform may entice consumers to consume more, potentially turning deadweight loss—music that is valued above zero but below its purchase price, and hence is not consumed—into surplus (e.g., see also Waldfogel 2012). Another possibility is that consumers merely shift consumption to other platforms: the songs may be different, but the time spent listening remains unchanged.

Second, music is a consumption good for which many consumers have a love of variety. We study the effects of streaming on the nature and magnitude of the music variety consumed. In the (digital) ownership model, when consumers purchase and download specific music titles, an additional variety is costly. However, in the streaming model, it is free. We measure how this price difference affects the breadth of consumed variety in terms of the number of distinct artists, songs, and genres consumed. Next, we measure how users reallocate their time listening: whether they concentrate on a few artists (e.g., superstars; Elberse 2008, Rosen 1981) or spread their listening across a wider set of artists.

Third, we investigate the impact of adopting a streaming technology on how consumers explore and discover new music. By reducing the costs of exploring the variety of music, streaming allows—according to one estimate—the average user to discover 27 new artists per month (Kissel 2015). To what extent do these new discoveries actually yield highly played songs? We investigate this issue empirically by considering how streaming affects repeat consumption for new titles in general, and for personal favorites among new titles.

We examine the effects of streaming by focusing on the moment of adoption of Spotify, currently the largest streaming provider serving 100 million customers in 60 countries (Spotify 2017a). We construct a unique panel data set capturing individual-level music consumption by using a third-party service that tracks consumers' platform choices and listening behavior across a wide set of platforms. We identify self-treated consumers who adopt streaming in our observation window, though they may continue to use other providers. Next, we match adopters to control users who do not adopt streaming. We also measure variety using a secondary data set with metalevel characteristics for more than 200,000 artists.

We identify the effects of adopting a streaming service on total music consumption on all platforms in the short (within two weeks), medium (up to six months), and long run (six to 12 months after adoption). We use a difference in differences (DiD) approach that controls for unobserved user-level and time-varying characteristics. Two forms of selection complicate our identification strategy. First, our data lack a randomized assignment of consumers into treatment and control conditions. We use a combination of consumer fixed effects and quasi-experimental methods to distinguish causal adoption effects from simple differences in the characteristics between adopters and nonadopters (e.g., Bronnenberg et al. 2010). Second, the demographic profile of sampled consumers may not be representative of the total population of potential adopters. Hence, we study local average treatment effects (LATEs) among those consumer segments in our sample that are likely to adopt streaming.

We find that the adoption of streaming leads to a long-run growth of 49% in overall music consumption across all platforms. A sizeable amount of consumption on streaming services comes at the expense of ownership platforms such as iTunes, Winamp, and Windows Media Player. Next, breadth of variety increases and concentration of listening behavior drops as consumers expand their listening over a larger assortment of artists, songs, and genres. Finally, streaming increases the rate at which consumers discover new music. Although consumers typically play a new song less after adopting streaming, the low-cost trial and wider selection of free additional music available on streaming services results in more plays for users' best discoveries. We examine several sources of heterogeneity in the local average treatment effect. For instance, relative to consumers whose listening histories contain many varieties, users with low variety prior to adoption typically experience larger adoption effects on variety and discovery of new music. We repeat our analysis with different variable operationalizations, functional forms, long-run effects, and definitions of our sample. The results from this and other robustness checks indicate broad agreement with our reported results.

Taken together, our results demonstrate a significant long-run shift in music consumption toward more plays, variety, and new music discovery. Hence, if we believe our sample is representative of the population of streaming adopters, the shift from ownership to streaming potentially levels the playing field to the benefit of smaller producers, e.g., indie artists or labels.

Section 2 provides the theoretical background in terms of consumer behavior. Sections 3 and 4 discuss the data and methodology. Section 5 presents the results, and Section 6 discusses the implications for producers and consumers of music. Section 7 concludes.

2. Variety in the Music Entertainment Industry

The music industry has been studied in economics (Adler 1985, Cameron and Collins 1997, Chung and Cox 1994, Rosen 1981), marketing (Chung et al. 2009, Holbrook and Hirschman 1982, Lacher and Mizerski 1994), law (Zentner 2006), and sociology (Lopes 1992). Because music variety is at the heart of consumer welfare, unsurprisingly, a central issue in the literature is the limits on variety consumption in demand and supply.

On the demand side, variety in music can serve two purposes. First, it can cater to consumers with idiosyncratic tastes (Crain and Tollison 2002). In this setting, more variety meets the tastes of more consumers and enhances welfare along the extensive consumer margin. Alternatively, a broad selection in music can satisfy the demand for variety at the individual level (Adler 1985, Chung and Cox 1994, Kim et al. 2002, Ratner et al. 1999), creating welfare along the intensive consumer margin. Another literature stream does not consider its search cost (see, e.g., Elberse 2008), but the acquisition cost of purchasing quantity versus variety (see, e.g., Bronnenberg 2015). This literature suggests that the costs of a marginal variety lead to limited demand for variety.

On the supply side, one of the earliest empirical observations concerning variety in the music industry is that a relatively small number of artists commands a large share of the revenue, i.e., that the music industry is characterized by limits to the supply of variety. Rosen (1981) presents a theory that proposes that "superstars" emerge from two conditions. One is that artists' rewards are convex in talent. This occurs when consumers view variety and quality as substitutes and are willing to pay more for a single top performance than for several good but intermediate ones. Convexity turns small talent differences into large payoff differences. The second condition is that the marginal cost of producing a consumer experience, e.g., the cost associated with producing an additional audio file, is low. This creates a scale economy allowing few sellers to serve many consumers.

An alternative view is that concentration arises—even among equally talented artists—from differences in search cost or complementary "consumption capital." For instance, Adler (1985) and Chung and Cox (1994) view music as an experience good. Rather than requiring a distribution of talent (and a magnified distribution of rewards), superstardom can emerge from imitation behavior by fans who have incomplete information and choose popular artists to minimize search costs. Adler (1985, p. 208) discusses this mechanism in the context of accumulating consumption capital that is produced by listening to music and "discussing it with other persons who know about it." This produces a reinforcing spillover among consumers and selects a handful of lucky performers to become stars.

Empirically, the consumption capital explanation has gained some support. Lacher and Mizerski (1994) find that consumers are likely to purchase music more by its ability to create an absorbing experience than by liking alone. Using a measure of voice quality, Hamlen (1991) documents that small differences in talent do not lead to excessive differences in rewards.

Entry of streaming providers in a market dominated by ownership models has affected the three different mechanisms of variety reduction mentioned above. First, and central to our paper, the ownership model charges a fee per variety, i.e., per song, whereas a streaming provider charges a subscription fee for the entire catalogue, thus dramatically lowering the acquisition cost of variety. Indeed, variety on a streaming provider is free of charge (although some search cost may remain), and this paper seeks to exploit this shift in costs to measure its impact on consumption and demand.

Second, there are additional effects on the supply of variety that may have secondary effects on variety consumption. Viewed through the lens of limited entry from convex rewards (e.g., Rosen 1981), a streaming provider changes an artist's reward structure to a fee per play that does not depend on the quality or overall popularity of an artist. Thus, relative to a world where artists that are more popular command higher prices, the streaming reward schedule is less convex in popularity. Additionally, streaming provides low-cost information about artists and measures of consumption capital to consumers (e.g., by providing consumption information in the form of playlists).

3. Data

We now describe the data that we have collected to shed light on how streaming changes the quantity and variety of music consumption and the discovery of new music.

3.1. Institutional Background

Recently, the music industry has witnessed a marked increase in the number of interactive streaming providers, which are the focus of this study.[2] At present there are over 20 providers offering comparable services in terms of variety and price.[3] The largest of these, Spotify, has 40 million paying subscribers in 60 countries (Spotify 2017a); subscribers have unlimited access to a library of over 30 million songs in exchange for about $10 per month. More than 60 million users are on Spotify's free membership plan. They have less control over what they can listen to, and advertisers pay Spotify to expose free users to commercials. Importantly, streaming providers pay royalties to copyright holders based on the number of times a song is streamed.

3.2. Sample

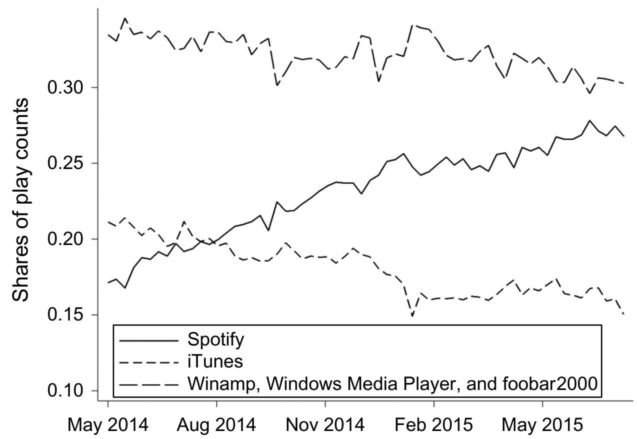
To construct our data, we used a third-party music recommendation service that wishes to remain anonymous (henceforth, "the service"). The service builds detailed user profiles by recording users' listening histories across multiple platforms. Consumers join it to receive music recommendations based on consumption across all their music sources rather than those based on only one platform. The service supports more than 100 devices and clients, giving a comprehensive picture of music consumption, whether offline or online, mobile or desktop.[4]

We sampled from the service's user base by repeatedly visiting its website in short intervals between April 22 and April 29, 2014 (Oestreicher-Singer and Zalmanson 2013). Our web scrapers collected the usernames 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 randomly selected sample of 5,003 users,[5] and retrieved their music consumption histories for a period of 2.5 years (January 6, 2013, until August 1, 2015). For each individual listener, we collected a unique username, a time stamp, and artist and song names. Platform choices were not directly observable via the service's API. Instead, we scraped users' platform choices from their profile pages between May 1, 2014, and August 1, 2015. Because of interruptions of our scrapers for technical reasons, we recorded consumption on 431 out of 458 days. Missing days were not systematic. We observed 123 million plays for 4,033 active users. Data on the remaining 970 users were not available, either because these users changed their privacy settings or cancelled their service accounts during our observation period. We defined an initialization sample to identify when users listened to content for the first time and to match adopters with users that did not adopt streaming based on their listening histories. This sample runs from January 6, 2013, to May 28, 2014, and includes four weeks in which we recorded users' platform choices via web scraping. We use the remaining 62 weeks of our data as the estimation sample, covering the period between May 29, 2014, and August 1, 2015.

We identify unique artists and songs using an algorithm described in Online Appendix A, so that "The Beatles" and "beatles" are counted as the same artist. Some songs have unrealistically short song lengths; also, users may skip songs after listening for a few seconds. We removed all songs shorter than 30 seconds and songs that were skipped before half of the song was finished (in total, 7.2% of all plays). We retained 114 million plays for 4,033 users. In our international sample, the most played artists are Lana Del Rey, Taylor Swift, and Madonna; the main genres are rock, pop, and metal;[6] and average daily music consumption is about three hours, close to the U.S. average of 3.6 hours for total music consumption (25 hours per week; see Nielsen 2014). The top platforms are Spotify and iTunes, with market shares of 22.8% and 18.3%, respectively, in terms of the total number of plays. Other platforms are mainly used to play locally stored MP3 files or listen to CDs on the computer, such as Winamp (12.5%), Windows Media Player (9.7%), and foobar2000 (10.1%; henceforth, collectively referred to as WWF). We group all remaining platforms in an "other" category[7] because they cannot be classified as streaming-based or ownership-based platforms (e.g., capturing consumption on a variety of websites), because they are noninteractive (see Endnote 2), or because their market share in our sample is negligible.[8]

In Figure 1, we plot the market shares for the major platforms in terms of play counts in our data. The figure shows that usage of Spotify grew steadily while that of iTunes and the remaining platforms declined. Hence, our aggregate data seem to confirm reports that Spotify is encroaching on the market shares of ownership-based platforms.

Figure 1. Market Shares (Play Counts) in the Raw Data



3.3. Identifying Adoption of Music Streaming

We consider Spotify as the only streaming platform because other streaming providers in our sample have negligible market shares. Because of potential left-truncation problems, we classify users as Spotify adopters only if we observe them not use Spotify for at least 45 days since the beginning of their recorded usage. Furthermore, we require potential adopters to be active on at least one of the major ownership-based platforms (i.e., iTunes or WWF). This procedure gives us 507 users that adopt Spotify in our 62-week estimation sample.[9] These adopters may continue to consume music on other platforms. A total of 1,471 users never adopt Spotify. We disregard 1,135 users who use Spotify before our 45-day cutoff, 714 users who never use a major ownership-based platform, and 206 light users who never listen in the initialization period or listen for less than 10 weeks in the estimation period.

How close is our sample to the population of Spotify or other potential streaming subscribers? Spotify (2016) reports that their adopters are largely millennials (born 1980–2000) and more likely to be male. These characteristics are mirrored in our sample (mean age is 22.44 years, and 74% of the users are male).[10] Furthermore, our data collection on streaming adoption (2014–2015) mostly takes place in markets where Spotify has been active since 2008, and hence largely captures early/late majority adopter segments, which probably have weaker tastes for variety than the innovators and early adopter segments in 2008–2014. For example, Spotify (2017b) reports daily music consumption for cross-platform users on an ad-based/free subscription at 148 minutes. Average daily consumption in our sample is one hour,[11] but this averages over both free and premium adopters, as well as adopters that listen on only one platform. Unfortunately, other consumption data on the population of Spotify adopters are not available to us. Hence, we cannot reliably rule out the possibility that our sample of adopters may not be representative of the larger population of potential streaming adopters.

Crucial to our investigation is that no major changes occurred to Spotify during our observation period. Had such changes occurred, then estimated treatment effects may have been confounded. We provide an overview of all significant changes to Spotify in Online Appendix B. In the first year of our estimation sample (between May 29, 2014, and April 27, 2015), major developments were confined to launching in new markets (Canada and Brazil), client updates (e.g., for Windows Phone), partnerships with hardware manufacturers (e.g., Sony's PlayStation), and the introduction of a small number of new playlists ("Dinner," "Sleep," "Folk Americana"). Furthermore, the supply of music remained stable (except for the removal of Taylor Swift and the introduction of John Lennon and Rammstein). Important changes to Spotify's recommendation algorithm were introduced in the last 16 weeks of our estimation period, such as "perfect playlists at your fingertips" (April 28, 2015), playlists for workouts ("Spotify Running") and an improved starting page (both May 20, 2015), and personalized playlists with previously unheard content ("Discover Weekly," July 20, 2015). In Section 5.5, we explain how we verify the robustness of our findings with regard to these innovations. Spotify's main competitor, Apple Music, was launched in the last four weeks of our data (on June 30, 2015).

3.4. Measuring Quantity, Variety, and Discovery

3.4.1. Quantity. The unit of analysis is the user-week: we measure the number of songs each user listens to in a week (i.e., the weekly play count) on any platform.

3.4.2. Variety. Measuring product variety is complex and difficult (Alexander 1997). We partially characterize variety by deriving a set of metrics in two categories: breadth (e.g., Hoch et al. 1999) and concentration (e.g., Elberse 2008). We specify these metrics in a multiattribute space (i.e., songs, artists, and an artist's genre), facilitating interpretation in favor of collapsed and hence more abstract metrics (e.g., entropy; van Herpen and Pieters 2002).

The first set of variety metrics relates to the *breadth of variety* consumed by users. We measure breadth by counting the number of distinct artists, songs, and genres consumed. The second set of metrics relates to the *concentration of variety*. We measure a user's tendency to listen predominantly to common favorites (superstars) as the listening share of top 20, top 100, and top 500 artists. Inclusion in the set of top artists is determined by ranking artists in terms of total plays by geographic region in a rolling window of one year, lagged by four weeks to avoid simultaneity bias ($t-55, \ldots, t-4$; see, e.g., Zentner et al. 2013). To measure how users allocate listening to their own personal favorites, we calculate the Herfindahl index of a user's weekly listening history for artists, songs, and genres.

3.4.3. Discovery. We measure the consumption share of new music. We define newness as a flow metric: an artist, song, or genre can be new to the user only in the first week of consumption. To operationalize newness, we use a long history of listening behavior starting on January 6, 2013. Newness is therefore specific to an individual's listening history, irrespective of the release date of the music.

As a proxy for how new music is valued, we measure repeat consumption share for new discoveries. To proxy for the value of the "best" tracks among these new discoveries, we calculate the ratio of top new plays to top overall plays (either new or known) over a rolling eight-week window.

We explain the operationalization of all metrics in Table 1 and provide summary statistics in Table 2. In Table C1 in Online Appendix C, we compare treatment and control groups on all measures. The differences reported in this table are not meant to represent an adoption effect, because they ignore consumer heterogeneity, common time trends, and any differences in adoption effects in the short and long run.

3.4.4. Do Our Measures Reflect Demand or Supply? A central concern to our investigation is whether the effects of adopting Spotify are due to changes in demand or supply on streaming. For example, if certain features restrict consumer choice or forbid it altogether, as is the case with the Internet radio provider Pandora, the results of our analysis will not be informative about whether the effects are due to consumer choice or the variety supplied. However, Spotify is an interactive streaming provider, and users can request content suggestions as a deliberate choice. Even if users were to receive recommendations (e.g., by listening to curated playlists), they could simply skip less preferred songs.[12] In other words, we assume that consumers are free to make their own consumption choices. Hence, we attribute changes in music consumption to shifts in demand, because it is unnecessary and unlikely that consumers listen to content on Spotify against their will.

Second, a related concern is that differences in assortment (e.g., size, or indie versus major) between ownership-based and streaming platforms may drive our effects. The music catalogues of both iTunes and Spotify feature around 30 million songs and are thus comparable (Mitroff and Blanco 2015). Furthermore, windowing strategies that released popular content on Spotify later than on iTunes were introduced only

Table 1. Variable Operationalization

Dimension	Operationalization
(1) *Quantity*	• Log number of song plays
(2) *Variety*	
Breadth	• Log number of unique artists, songs, and genres listened to
Concentration in common favorites (superstars)	• Number of unique artists in the top 20, top 100, and top 500 in a user's geographic region[a] (ranked according to a rolling window of a year, lagged by four weeks; $t-55, \ldots, t-4$), divided by the number of unique artists listened to over the same time period
Concentration in personal favorites	• Herfindahl index (sum of squared listening shares), computed over a user's weekly plays of artists, songs, and genres
(3) *Discovery*	
New content consumption	• Number of distinct new artists, songs, and genres listened to by a user for the first time,[b] divided by the total number of distinct artists, songs, and genres listened to
Repeat consumption	• Number of unique new artists, songs, and genres played more than once, divided by the total number of unique new artists, songs, and genres listened to
Best discoveries	• Amount of plays of the top 1 new artist, song, and genre in an 8-week period subsequent to discovery ($t+1, \ldots, t+8$) ranked in order of plays, divided by the amount of plays of the overall (not necessarily new) top 1 artist, song, and genre over the same time period

Note. All variables are computed at the user-week level.

[a]Regions are the European Union, South America, United States and Canada, and others.

[b]This refers to content consumed in a user's first week of consumption, based on the users' listening history on the service up to January 6, 2013.

after our observation period. We acknowledge that, typically, there are more songs available for purchase than for streaming (e.g., CDs bought at a local concert). However, an assortment of millions of songs on streaming platforms is not likely to constrain choice in a meaningful way.

4. Method

4.1. Identification Strategy

Our objective is to identify LATEs of adopting Spotify on quantity, variety, and discovery. We face two major challenges. First, our data generation process lacks a randomized assignment of consumers into treatment and control conditions. A simple approach is to estimate a DiD model, which removes any persistent linear, additive consumer-specific effects that may, if ignored, introduce endogeneity due to self-selection into adopting Spotify. In addition, for our main results, we assume that we can control for the unobserved need for variety in absence of streaming by conditioning on a rich set of observed characteristics (for a similar approach, see Bronnenberg et al. 2010); that is, we use a quasi-experimental matching procedure in which we match adopters with similar nonadopters based on a propensity score constructed from variables that capture users' demand for variety.

Second, treated and control consumers may exit our sample at different moments (e.g., because they stop using the service). To achieve comparability in terms of market trends, we seek pairs of users with similar beginning and end points in our observation period. We combine all criteria by selecting pairs that minimize the Mahalanobis distance between treated and potential control users based on the propensity score and beginning and end points in our observation window. We now explain in detail each step of our identification strategy.

4.2. Comparison of Treated and Control Groups

As a first step, we assess whether nonadopters display the same listening behavior as adopters prior to adoption. To this end, we compute our measures of listening behavior during a 12-week period before the start of our estimation sample. Table C2 in Online Appendix C shows that, on average, adopters and nonadopters differ significantly on a set of key behavioral and demographic variables. Adopters listen to more songs (49.95 plays versus 42.27 plays, $p < 0.001$) on more platforms (3.17 versus 2.87, $p < 0.001$) and tend to consume more mainstream music (share of plays to top 100 artists, 0.24 versus 0.15, $p < 0.001$) than nonadopters. Adopters consume fewer artists (3.32 versus 4.00, $p = 0.001$) and discover less new content (share of plays 0.16 versus 0.22, $p < 0.001$). Adopters are also younger (22.44 years versus 24.13 years, $p < 0.001$). Because users in the treatment and control groups differ on several behavioral and demographic characteristics, it is likely that mere differences in the composition of these groups could explain differences in their behavior. We therefore use quasi-experimental methods to deal with selection effects.

4.3. Propensity Score Estimation and Matching

Self-selection may arise because of differences in taste for variety. For example, variety-loving consumers may adopt streaming because content discovery on ownership-based models is too expensive for them. In principle, this type of selection can be accounted for by using individual fixed effects, which we indeed include in all of our models. However, fixed effects impose a

Table 2. Summary Statistics

	N	Mean	SD	Min.	Max.
User characteristics					
Gender (female = 1)	1,978	0.24	0.43	0.00	1.00
Age	1,978	23.70	6.23	11.00	70.00
European Union (dummy)	1,978	0.32	0.47	0.00	1.00
South America (dummy)	1,978	0.21	0.41	0.00	1.00
United States/Canada (dummy)	1,978	0.10	0.30	0.00	1.00
Other geographic region (dummy)	1,978	0.37	0.48	0.00	1.00
Quantity of consumption					
Play counts on all platforms	122,636	188.89	269.43	0.00	3,862.00
Play counts on Spotify	122,636	11.19	69.38	0.00	2,439.00
Play counts on iTunes	122,636	48.43	148.59	0.00	3,857.00
Play counts on Winamp, Windows Media Player, and foobar2000	122,636	94.32	211.80	0.00	2,945.00
Play counts on other platforms	122,636	34.96	110.17	0.00	2,432.00
Breadth of variety					
Number of unique artists	97,924	36.64	50.92	1.00	882.00
Number of unique songs	97,924	150.44	169.79	1.00	2,603.00
Number of unique genres	97,924	14.00	13.49	1.00	209.00
Concentration of variety					
Top 20 artists (share of unique artists)	97,924	0.04	0.09	0.00	1.00
Top 100 artists (share of unique artists)	97,924	0.12	0.17	0.00	1.00
Top 500 artists (share of unique artists)	97,924	0.29	0.25	0.00	1.00
Artist concentration (Herf.)	97,924	0.21	0.23	0.00	1.00
Song concentration (Herf.)	97,924	0.05	0.11	0.00	1.00
Genre concentration (Herf.)	97,924	0.35	0.24	0.03	1.00
Discovery of new content					
New artists (share of unique artists)	97,924	0.20	0.22	0.00	1.00
New songs (share of unique songs)	97,924	0.37	0.27	0.00	1.00
New genres (share of unique genres)	97,924	0.05	0.09	0.00	1.00
New artists played more than once (share of unique new artists)	75,116	0.59	0.37	0.00	1.00
New songs played more than once (share of unique new songs)	91,579	0.22	0.25	0.00	1.00
New genres played more than once (share of unique new genres)	33,617	0.56	0.45	0.00	1.00
Top 1 new artist to overall top 1 artist (share of plays)	80,061	0.22	0.33	0.00	1.00
Top 1 new song to overall top 1 song (share of plays)	77,239	0.55	0.41	0.00	1.00
Top 1 new genre to overall top 1 genre (share of plays)	83,445	0.01	0.07	0.00	1.00

Notes. Summary statistics are calculated on an unmatched sample of 507 adopters and 1,471 nonadopters observed over 62 weeks starting May 29, 2014. The unit of analysis is an individual user for user characteristics and the user-week for the remaining variables. Herf., Herfindahl index.

linear functional form on the baseline, whereas matching allows for more flexibility (such as potential interactions between the baseline and effects in the short and long run). In addition, matching removes from the sample users that are likely to never adopt Spotify, such that the control group is comprised of users that resemble adopters in every way except for their adoption. Hence, we combine matching and fixed effects to make the comparison between adopters and nonadopters as close as possible.[13] Specifically, we pair each adopter of Spotify with a comparable consumer that is very likely to adopt but "randomly" did not do so in our 62-week observation period.

We execute our matching procedure in three steps: First, we estimate each household's adoption propensity as a function of observed variables (e.g., Rosenbaum and Rubin 1983)

$$\Pr(\text{adopt}_i = 1) = \Pr(\beta_0 + Z_i\gamma + \eta_i > 0), \quad (1)$$

where Z_i is a vector of observed household-specific characteristics. The covariates entering Z_i describe a user's listening behavior (e.g., quantity, variety consumption, and discovery) during the 12 weeks prior to our estimation period starting May 29, 2014. The relevant descriptive statistics are in Table C2 in Online Appendix C. Our matching estimator is static: it uses the same preadoption observation window to predict adoption for early versus late adopters. We believe our assumption of more or less stable predictors is tenable, as adoption occurs in a relatively narrow window of about a year. We assume that the η_i are independent and identically distributed random variables with a type I extreme value distribution, making the probability in Equation (1) a binary logit model.

Second, for further comparability of the treatment and control samples, we ensure that users in a matched treated–control pair have the same observation window. We observe most users during our complete observation period, but about 10% of users stop using the service and exit the sample. If treated users were matched with control users in a different observation window, or vice versa, any estimated treatment effect would not be clearly attributable to adoption but could arise from timing differences across consumers in recording music consumption on the service. Thus, to ensure we are using the same observation window for matched treated–control pairs, we match users by their first and last observed periods of listening in our sample.

In the third stage, we construct a three-dimensional distance metric of the composite of propensity score and the first and last periods of consumption. We compute the Mahalanobis distance for each treated and control user pair and use the one-nearest-neighbor algorithm to match users that are closest together. We drop control users outside the region of common support, defined as the overlap in propensity scores between treated and control users (e.g., Gensler et al. 2012). Some of the matches that we obtain are potentially nonunique; i.e., there are instances where the same nonadopter is the closest match for more than one adopter. We select unique matches sequentially, in order of closeness of their Mahalanobis distances. To ensure a sufficient matching quality, while avoiding the need to match with replacement, we match an adopter to her kth best match. This strikes a balance between the number of adopters in our sample (increasing in k) and matching quality (decreasing in k). In preparing our final data set, we set k to 3 and retain those user periods in which treatment–control observations overlap.

The matching procedure yields 448 treated and 448 matched control users. In Table C3 in Online Appendix C, we report the results of the logit propensity score model. The directions of the effects have face validity: ceteris paribus, consumers with higher average plays, more superstar consumption, and more platforms used are more likely to adopt Spotify. In turn, older users with more concentrated listening and users who already have access to and listen to new music are less likely to adopt Spotify. Notably, users from South America and the United States/Canada have higher adoption probabilities than users from other geographic regions, as Spotify launched in Brazil and Canada during our observation period. The hit rate of the model is 66%, and the McFadden pseudo-R^2 is 0.092.

After matching, both consumer groups are indistinguishable in terms of their adoption propensities (see Figure 2), observation windows (see Figure C2 in Online Appendix C), and observed demographics and preadoption behavior. In the bottom part of Table C2 in Online Appendix C, we report presample summary statistics for the matched control sample, which are now very close to those of the matched treatment group.

4.4. Difference in Differences

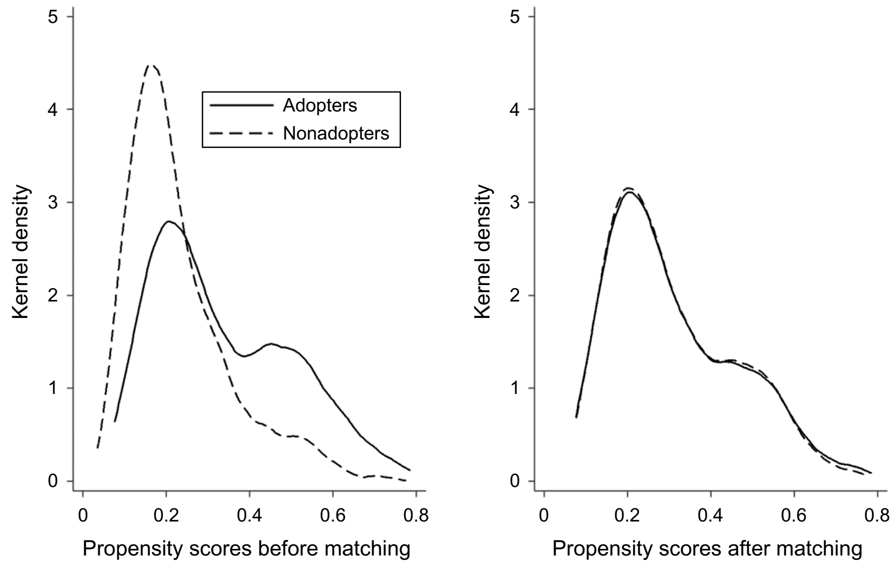
We use a DiD approach to estimate the effect of adopting streaming on our outcome measures. We compare the outcome measures of adopters before and after their adoption with those of the matched nonadopters. We also investigate how long changes in consumption, variety, and discovery last. We estimate models of the following type:

$$\begin{aligned} Y_{it} = {} & \alpha_i + \gamma_t + \beta^{\mathrm{ST}} \cdot I(0 \le \text{weeks\_since\_adoption}_{it} \le 1) \\ & + \beta^{\mathrm{MT}} \cdot I(2 \le \text{weeks\_since\_adoption}_{it} \le 24) \\ & + \beta^{\mathrm{LT}} \cdot I(\text{weeks\_since\_adoption}_{it} \ge 25) + \varepsilon_{it}, \end{aligned} \tag{2}$$

where Y_{it} is the dependent variable, and the indicator variables $I(\text{weeks\_since\_adoption}_{it})$ are 1 if the number of weeks since adoption for consumer i is within the indicated range, in week t. Furthermore, α_i is a consumer-level fixed effect, γ_t is a week-level fixed effect, and ε_{it} is the error. This two-way fixed effects specification controls for time-invariant consumer characteristics, such as overall liking of music, as well as common time trends and week-to-week fluctuations. An important identifying assumption is that the ε_{it} are orthogonal to the indicator variables $I(\text{weeks\_since\_adoption}_{it})$. We distinguish between treatment effects in the short run (within the first two weeks of adoption, β^{ST}), medium run (between 2 and 24 weeks after adoption, β^{MT}), and long run (25 weeks and after, β^{LT}). We use robust standard errors clustered at the user level to account for any serial correlation (Bertrand et al. 2004).

The DiD approach relies on the assumption of parallel pretreatment trends. To test whether it holds, we carry out so-called "placebo" treatment tests using preadoption data (and matching the observation window in the control sample). In particular, we define a placebo "treatment" at the midpoint of a user's pretreatment data. Next, we estimate a DiD model for all dependent variables listed in Table 1. For each of the 22 combinations of variables and dimensions, we fail to reject the null hypothesis of no treatment effect for

Figure 2. Distribution of Propensity Scores Before and After Matching



placebo treatments (see Online Appendix C, Table C4). This supports the idea that the pretreatment trends are statistically equivalent across both user groups.

Taken together, we combine a propensity score matching with a DiD approach. Our propensity score model selects nonadopting consumers who are like adopters, except for not subscribing to music streaming. Because we include a rich set of behavioral measures of preexisting consumer tastes for variety and discovery in our propensity score model, we assume that the unobserved components in the propensity Equation (1), i.e., the η_i, are independent of the unobserved components of our regression model (2), i.e., the ε_{it}.

Given our method, we interpret the reported effects from the DiD regressions as the average treatment effect on the adopters (compliers), i.e., as LATEs. In this context, we highlight one additional type of analysis that our data permit us to run. Because adoption time differs across adopters (see Online Appendix C, Figure C1), we can estimate a treatment effect using only within-adopter variation, with late adopters acting as a control for early adopters. This approach also accounts for possible selection on unobservables shared by late and early adopters. We report on this analysis as a robustness check, after discussing the main results.

4.5. Heterogeneity in Treatment Effects

The impact of streaming may depend on individual characteristics and circumstances. In this section, we consider how three potential user-level moderators affect the discovery of new music: (1) preexisting consumption of variety, (2) user age, and (3) free versus premium subscription. There is little existing research about the sign of these effects, and our discussion of them is largely exploratory.

First, the effect of adopting streaming on new music discovery may differ across users who originally consume limited versus extensive variety. The sign of this moderation is theoretically ambiguous. On the one hand, preexisting consumption of variety may be low because its cost suppresses demand. To such consumers, making additional variety free putatively has a big effect on variety seeking and discovery. To consumers who already listen to a lot of variety before adopting, in turn, the adoption effect may be less due to a constraint on extra listening time. On the other hand, the moderation can be oppositely signed. Consumers may listen to few varieties prior to adoption because they have a limited taste for it. To them, adopting Spotify likely results in a smaller effect on new music discovery compared to high-variety consumers. We study this moderation effect empirically by estimating the interaction of adoption and users' preexisting taste for variety, inversely measured as artist concentration in each user's listening history in the 12 weeks before the start of the estimation period. In particular, we use the Herfindahl index computed as the sum of squared artist shares for each user (median = 0.034).

Second, we are interested in the moderating effect of age. This interaction effect is also theoretically indeterminate. On the one hand, taste for certain music genres may develop with age, and, hence, young users' exploration of new varieties of music may be limited, even when variety is free. On the other hand, younger consumers are more income constrained and would therefore benefit the most from new content being free at the margin. To empirically determine which effect dominates, we estimate the moderating effect of each user's age (median = 22 years) on adoption.

Last, Spotify's free plan users may have less control than paying subscribers over what they consume;

hence, they may see fewer discoveries and varieties compared to premium subscribers. To test for this effect, we estimate the interaction of Spotify adoption with whether users are on the premium or ad-based (free) subscription plan of Spotify (median = 0.561 ads an hour).[14]

We incorporate each of these terms as moderators of an adoption effect. To avoid cluttering, we do not distinguish between short-, medium-, and long-run effects, but use a single treatment effect ($adoption_{it}$), which equals one on and after the week of adoption and zero otherwise. To enhance interpretability, we use effect coding for the moderators (e.g., $Herf_i = 1$ if the user has an above-median Herfindahl score, and $Herf_i = -1$ if the score is below the median). The interpretation of δ is the average adoption effect, while ϕ is the deviation from that effect for an above-median (versus below-median) consumer. This is specified as follows:

$$Y_{it} = \alpha_i + \gamma_t + \delta \cdot \text{adoption}_{it} + \phi_1(\text{adoption}_{it} \cdot \text{Herf}_i) + \phi_2(\text{adoption}_{it} \cdot \text{Age}_i) + \phi_3(\text{adoption}_{it} \cdot \text{Free}_i) + \varepsilon_{it}. \quad (3)$$

5. Results

5.1. Consumption Growth and Displacement

The first questions we address are whether adopting Spotify leads to extra music consumption and how long these effects last. In Table 3, column (1), our dependent variable is the log total play count consumed across all platforms by a given user in a given week. In the week of adoption and the week after, the number of plays grows by 132% (= exp(0.84) − 1). Total consumption is 63% higher (= exp(0.49) − 1) in the medium run, from two weeks until 24 weeks after Spotify adoption. Even more than 25 weeks (nearly six months) after adoption, overall consumption is still 49% higher (= exp(0.40) − 1) than before adopting Spotify.

To what extent does Spotify adoption displace consumption on iTunes and other ownership-based platforms? In column (2) of Table 3, we see that iTunes consumption drops 21% in the first two weeks after Spotify adoption, drops 27% in the following 22 weeks, and is 28% lower about six months after Spotify adoption. Column (3) shows that consumption on WWF falls 22% in the first two weeks after Spotify adoption, 34% in the following 22 weeks, and 37% after six months. Both displacement effects grow over time. The impact on consumption on all remaining platforms is not significant in the first 24 weeks; after six months there is a 24% drop in column (4).

To sum up, first, Spotify adoption leads to strong consumption growth that persists beyond 24 weeks after the moment of adoption.[15] Second, and in accordance with Figure 1, we see that Spotify increasingly displaces consumption from iTunes and WWF.[16]

5.2. Variety Consumption

5.2.1. Breadth of Variety. We next investigate the effect of Spotify adoption on the variety of music consumption. Recall that adopting Spotify lowers the monetary cost of the marginal variety to zero. Hence, to the extent that this cost limits demand for variety, we expect users to consume more variety after adoption.

Table 3. Adoption of Streaming Increases Total Consumption Yet Cannibalizes from iTunes and Other Platforms

	(1) Log play counts on all platforms	(2) Log play counts on iTunes	(3) Log play counts on Winamp, Windows Media Player, and foobar2000	(4) Log play counts on other platforms
Short run (0–1)	0.84***	−0.23**	−0.25***	0.039
	(0.066)	(0.072)	(0.070)	(0.067)
Medium run (2–24)	0.49***	−0.31***	−0.42***	−0.10
	(0.070)	(0.080)	(0.083)	(0.078)
Long run (25+)	0.40***	−0.33**	−0.46***	−0.27*
	(0.096)	(0.11)	(0.12)	(0.11)
R-squared	0.52	0.75	0.73	0.64
F	82.4	6.97	13.7	12.8
p-value	0.000	0.000	0.000	0.000
Users	896	896	896	896
Observations	52,346	52,346	52,346	52,346

Notes. The table shows a regression with robust standard errors in parentheses. Estimates are calculated on a matched sample of 448 adopters and 448 nonadopters observed over 62 weeks starting May 29, 2014. User- and week-specific fixed effects are used, and the unit of analysis is the user-week. The dependent variable is the log number of songs heard by a panelist in a week (play count). The independent variables are indicators for a user's time since adoption of Spotify, defined as short run (within weeks 0 and 1), medium run (within weeks 2 and 24), and long run (weeks 25 and after). Robustness checks are described in Online Appendix D.

$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

Table 4. Adoption of Streaming Increases Breadth of Variety

	(1) Log number of unique artists	(2) Log number of unique songs	(3) Log number of unique genres
Short run (0–1)	0.48***	0.40***	0.36***
	(0.040)	(0.041)	(0.030)
Medium run (2–24)	0.27***	0.26***	0.21***
	(0.039)	(0.040)	(0.029)
Long run (25+)	0.28***	0.27***	0.20***
	(0.054)	(0.057)	(0.041)
R-squared	0.56	0.55	0.56
F	50.1	65.9	43.2
p-value	0.000	0.000	0.000
Users	896	896	896
Observations	46,688	46,688	46,688

Notes. The table shows a regression with robust standard errors in parentheses. Estimates are calculated on a matched sample of 448 adopters and 448 nonadopters observed over 62 weeks starting May 29, 2014. User- and week-specific fixed effects are used, and the unit of analysis is the user-week when there is at least one song played. The dependent variables are the log numbers of distinct artists, songs, and genres heard by a panelist in a week. The independent variables are indicators for a user's time since adoption of Spotify, defined as short run (within weeks 0 and 1), medium run (within weeks 2 and 24), and long run (weeks 25 and after). Complete results and robustness checks are described in Online Appendix D.

$^{***}p < 0.001$.

Table 4 presents results for the log total numbers of unique artists, songs, and genres.

We find a significant increase in all measures immediately following Spotify adoption. In the first two weeks after adoption, the number of unique artists heard increases by 62%, the number of unique songs increases by 49%, and the number of unique genres increases by 43%. The effect attenuates over time, but consumption remains above preadoption levels both statistically and substantively. For instance, 2–24 weeks after adoption, the number of unique artists consumed weekly is 31% higher than preadoption levels. Similarly, after 25 weeks and up to approximately a year after adoption, we measure a 32% increase in the number of unique artists consumed. We see similar patterns for the other measures. These results strongly point to a permanent increase in the breadth of music consumption.

5.2.2. Concentration of Variety. *Superstar consumption.* One avenue through which superstars can arise according to Section 2—even in the absence of talent differences—is that consumers are uncertain about quality and "economize" on learning and search costs (Adler 1985). We argue that Spotify lowers search and learning costs by, among others, letting users costlessly consume from their catalogue of over 30 million songs and recommending playlists to its users. Hence, we expect superstar consumption to decline after adoption.

In Table 5 (columns (1) and (2)), we empirically investigate the impact of Spotify adoption on the consumption of artists in the top 100 and top 500. (The results for the top 20 are in Table D4 in Online Appendix D.) Because consumption quantity and variety increase after Spotify adoption (see Tables 3 and 4), we measure superstar consumption as a *share* of unique varieties. We find that users shift their consumption out of the top artists. Column (1) reports that in the first two weeks following Spotify adoption, the consumption share of top 100 artists drops by 0.028 in terms of unique varieties, from a preadoption baseline of 0.17 (see Table D4 in Online Appendix D). This drop represents a substantial 16% loss in superstar consumption. In the medium run, the superstar consumption share is still 0.015 lower (9% less) than preadoption levels. In the long run, the effects are 0.012 lower (7% less), but only marginally significant. Still, the effects are substantial in magnitude. A similar pattern holds for the top 500 artists (column (2)), with significant long-run effects at $p < 0.01$.[17]

Concentration in personal favorites. In addition to the diminishing consumption share of common favorites (superstars), consumers may allocate less of their listening to their own personal favorites, e.g., their own (weekly) top artists. Columns (3) and (4) of Table 5 investigate this using the Herfindahl index as a measure of concentration. The results show a clear decrease in concentration following Spotify adoption. For example, the Herfindahl index with respect to artists (column (3)) decreases from its preadoption baseline of 0.20 by 0.062 in the short run, 0.032 in the medium run, and 0.035 in the long run. The pattern also holds for songs (and genres; see Table D4 in Online Appendix D). These results show that music consumption fragments across a wider set of varieties as consumers allocate a smaller fraction of their total time to the top varieties each week. This shift in consumer behavior is persistent: the effect still holds in the long run.

5.3. Discovery

5.3.1. Consumption of New Content. If Spotify lowers the search cost for new music, we would expect that users discover more varieties, i.e., songs they have not consumed previously. We measure the share of listening of new varieties as a function of the number of unique varieties consumed each week.

Table 6 (columns (1) and (2)) shows the results for artists and songs. (The genre results are provided in Table D5 in Online Appendix D.) The adoption of Spotify results in a marked increase in the share of new content. For example, the variety share of new artists (column (1)) increases substantially from its preadoption level of 0.15 by 0.14 in the short run; it is 0.045 higher in the medium run, and still 0.032 higher in the

Table 5. Adoption of Streaming Leads to a Drop in Concentration

	(1) Top 100 artists (share of unique artists)	(2) Top 500 artists (share of unique artists)	(3) Artist concentration (Herf.)	(4) Song concentration (Herf.)
Short run (0–1)	-0.028^{***}	-0.042^{***}	-0.062^{***}	-0.022^{***}
	(0.0052)	(0.0066)	(0.0069)	(0.0031)
Medium run (2–24)	-0.015^{**}	-0.026^{***}	-0.032^{***}	-0.017^{***}
	(0.0048)	(0.0061)	(0.0064)	(0.0026)
Long run (25+)	-0.012^{+}	-0.024^{**}	-0.035^{***}	-0.016^{***}
	(0.0063)	(0.0084)	(0.0089)	(0.0040)
R-squared	0.57	0.63	0.35	0.22
F	6.94	14.5	13.8	6.05
p-value	0.000	0.000	0.000	0.000
Users	896	896	896	896
Observations	46,688	46,688	46,688	46,688

Notes. The table shows a regression with robust standard errors in parentheses. Estimates are calculated on a matched sample of 448 adopters and 448 nonadopters observed over 62 weeks starting May 29, 2014. User- and week-specific fixed effects are used, and the unit of analysis is the user-week when there is at least one song played. The dependent variables are (a) superstar consumption, defined as the number of unique popular artists in a user's geographic region (ranked by play counts over a rolling period of 52 weeks and lagged by four weeks to avoid simultaneity bias) divided by the number of unique artists, and (b) the concentration of a user's weekly play counts to artists and songs (measured using the Herfindahl index). The independent variables are indicators for a user's time since adoption of Spotify, defined as short run (within weeks 0 and 1), medium run (within weeks 2 and 24), and long run (weeks 25 and after). Complete results and robustness checks are described in Online Appendix D. Herf., Herfindahl index.

$^{+}p < 0.10$; $^{**}p < 0.01$; $^{***}p < 0.001$.

Table 6. Adoption of Streaming Leads to More Discovery That Is Repeat-Consumed Less on Average, Yet More at the Top

	(1) New artists	(2) New songs	(3) New artists played more than once	(4) New songs played more than once	(5) Top 1 new artist to overall top 1 artist	(6) Top 1 new song to overall top 1 song
Short run (0–1)	0.14^{***}	0.15^{***}	-0.093^{***}	-0.033^{***}	0.080^{***}	0.13^{***}
	(0.0097)	(0.010)	(0.013)	(0.0073)	(0.018)	(0.019)
Medium run (2–24)	0.045^{***}	0.047^{***}	-0.044^{***}	-0.019^{**}	0.016	0.056^{***}
	(0.0060)	(0.0077)	(0.011)	(0.0063)	(0.011)	(0.015)
Long run (25+)	0.032^{***}	0.022^{*}	-0.040^{**}	-0.016^{+}	0.00044	0.049^{*}
	(0.0078)	(0.0100)	(0.014)	(0.0089)	(0.016)	(0.020)
R-squared	0.36	0.43	0.33	0.36	0.22	0.22
F	25.2	18.3	2.82	4.01	17.1	10.0
p-value	0.000	0.000	0.000	0.000	0.000	0.000
Users	896	896	895	896	887	886
Observations	46,688	46,688	35,701	43,825	33,843	33,113

Notes. The table shows a regression model with robust standard errors in parentheses. Estimates are calculated on a matched sample of 448 adopters and 448 nonadopters observed over 62 weeks starting May 29, 2014. User- and week-specific fixed effects are used, and the unit of analysis is the user-week when there is at least one song played. The dependent variables are (a) the number of distinct new artists (songs) listened to divided by the total number of distinct artists (songs), (b) the number of distinct new artists (songs) played more than once divided by the total number of distinct new artists (songs) listened to, and (c) the share of plays to the top one new artist (song) in an eight-week window subsequent to discovery $(t+1, \ldots, t+8)$, ranked in order of plays, divided by the share of plays to the overall (not necessarily new) top one artist (song) over the same time period. For this metric, observations are excluded when the rolling eight-week window includes both preadoption and postadoption periods, and when there are fewer than eight weeks remaining at the end of each user's observation period. New artists and songs are defined by a user's first week of consumption on the service up to January 6, 2013. The independent variables are indicators for a user's time since adoption of Spotify, defined as short run (within weeks 0 and 1), medium run (within weeks 2 and 24), and long run (weeks 25 and after). Complete results and robustness checks are described in Online Appendix D.

$^{+}p < 0.10$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

long run. The pattern also holds for songs and genres. Thus, Spotify adoption accelerates the rate of new variety consumption.

5.3.2. Repeat Consumption and the Value of Discovery. A key question about the value of the variety expansion is whether the newly chosen variety has high match value. Empirically, the effect of adopting Spotify could go either way, depending on whether consumers have good information about their match value with artists, songs, and genres on the market a priori. If they do, then variety expansion from lower cost is likely limited to downward selection into new but less preferred content. However, if we believe that music is an experience good (Adler 1985), consumers' preconsumption valuations need not be the same as their ex post valuations (Rob and Waldfogel 2006). In short, consumers may not know their match value a priori. In this case, offering variety at a low cost leads to more experimentation and learning about new content. It may therefore lead to the discovery of new favorites or upward selection.

While a complete welfare analysis is beyond the scope of this paper, we perform two types of analyses to study the direction in which consumers expand variety. First, we use repeat listening for new content as a proxy for value, assuming that consumers will repeatedly consume content they like. In Table 6, columns (3) and (4), respectively, we measure the number of new (i.e., to the consumer) artists (songs) played more than once as a share of the total unique known artists (songs) consumed. (The genre results are in Table D5 in Online Appendix D.) The share of new artists consumed repeatedly drops by 0.093 directly following adoption. It is 0.044 lower in the medium run, and 0.040 lower in the long run compared to preadoption levels of 0.60. Similarly, column (4) shows that the share of repeatedly played new songs drops by 0.033 in the short run, 0.019 in the medium run, and 0.016 in the long run compared to preadoption levels of 0.22. Collectively, these results demonstrate that new content found on Spotify is subject to downward selection and that newly discovered music after Spotify adoption is, on average, of lower value.

Importantly, though, downward selection on average might simply be masking that consumers listen to larger quantities of new music to find new favorites. Consumers may be better off if their "best" discoveries are of high value. Hence, we investigate the effect of Spotify adoption on the match value of a user's *best* new artists, songs, and genres. In particular, we first construct the share of new music as the consumption that belongs to the user's top one artist, song, and genre that are new to the consumer in that week, where consumption is computed over the next eight weeks. Then, we take this share relative to the average consumption of the (not necessarily new) top one artist, song, and genre in that period. Hence, our measure is a ratio, similar to a top one lift for a given user, of her top new varieties to the best overall varieties. This measure is between zero and one, depending on whether the user's best discovery gets hardly any plays or whether it is the new leader even among past favorites.

Column (5) in Table 6 shows that the top new artist lift, i.e., the share of plays to the top new artist, immediately increases by 0.080 for adopters of Spotify, but is insignificant in the medium run and long run. Column (6) shows that the share of plays to the top new song increases by 0.13 in the short run and is 0.056 higher in the medium run. Twenty-five or more weeks after adoption, top new songs are consumed 0.049 more than the top songs discovered before Spotify adoption, suggesting that these new songs have higher match value. To conclude, columns (5) and (6) in Table 6 provide some evidence that the top new discoveries are consumed more (and hence provide more value to consumers) after than before the adoption of Spotify. The pattern also holds for genre consumption and is robust to alternative metrics (e.g., top lift computed over a user's five best discoveries, or consumption over a window of 12 weeks; see Table D5 in Online Appendix D). However, our estimate of the adoption effect on the value of top artist discoveries remains insignificant in the long run, regardless of how we operationalize this measure.

5.4. Heterogeneous Treatment Effects

There are sizeable effects of adopting Spotify on consumption, variety, and discovery. We now explore how these effects differ across adopters. As mentioned in Section 4.5, we focus on three particular moderators, using a median split on concentration of listening (high versus low), age (old versus young users; median is 22 years), and the frequency of 30 second (advertising) gaps in listening history as a proxy for free versus premium subscribers.

In Table 7, we investigate heterogeneous treatment effects for new music discovery; in Table C5 in Online Appendix C, we provide a full set of heterogeneous treatment effects for all outcome measures. Table 7 presents support across multiple measures that the Spotify adoption effect on new music discovery is stronger for users who (in the presample period) have more concentrated listening (i.e., with less variety). This is consistent with the notion that many users with low initial variety of consumption are constrained by the cost of, instead of low taste for, new discoveries. Additionally, older users in our sample discover more new content than younger users. For instance, the effect of adoption on new artist consumption (column (1)) is close to 0.036 for average adopters, but 41% larger (0.015) for users with above-median concentration, and 36% higher (0.013) for those with above-median age (older than 22). Last, the adoption effect on new music discovery appears largely similar for users on ad-based (free) and premium subscriptions.

Table 7. Assessing Potential Moderators of the Adoption Effect for New Music Discovery

	(1) New artists	(2) New songs	(3) New artists played more than once	(4) New songs played more than once	(5) Top 1 new artist to overall top 1 artist	(6) Top 1 new song to overall top 1 song
Adoption	0.036**	0.030*	−0.054**	−0.0075	0.018	0.054*
	(0.011)	(0.014)	(0.018)	(0.010)	(0.021)	(0.026)
× *Artist concentration*	0.015**	0.018*	−0.020*	−0.010+	0.020*	0.010
	(0.0055)	(0.0070)	(0.0094)	(0.0055)	(0.0097)	(0.013)
× *Age*	0.013*	0.019**	0.014	0.0015	0.0082	0.018
	(0.0054)	(0.0069)	(0.0092)	(0.0054)	(0.0098)	(0.013)
× *Free* (vs. premium)	−0.0069	−0.0083	0.0038	−0.0059	−0.018+	−0.014
	(0.0054)	(0.0070)	(0.0091)	(0.0055)	(0.0099)	(0.013)
R-squared	0.36	0.43	0.33	0.36	0.22	0.22
F	22.3	14.9	2.67	3.84	16.9	9.51
p-value	0.000	0.000	0.000	0.000	0.000	0.000
Users	896	896	895	896	887	886
Observations	46,688	46,688	35,701	43,825	33,843	33,113

Notes. The table shows a regression model with robust standard errors in parentheses. Estimates are calculated on a matched sample of 448 adopters and 448 nonadopters observed over 62 weeks starting May 29, 2014. User- and week-specific fixed effects are used, and the unit of analysis is the user-week when there is at least one song played. The dependent variables are (a) the number of distinct new artists (songs) listened to divided by the total number of distinct artists (songs), (b) the number of distinct new artists (songs) played more than once divided by the total number of distinct new artists (songs) listened to, and (c) the share of plays to the top one new artist (song) in an eight-week window subsequent to discovery ($t+1, \ldots, t+8$), ranked in order of plays, divided by the share of plays to the overall (not necessarily new) top one artist (song) over the same time period. For this metric, observations are excluded when the rolling eight-week window includes both preadoption and postadoption periods, and when there are fewer than eight weeks remaining at the end of each user's observation period. New artists and songs are defined by a user's first week of consumption on the service up to January 6, 2013. The independent variables are indicators for a user's adoption of Spotify and interaction effects with presample measures to capture heterogeneous treatment effects (median split, effect coding used so that the main adoption effect can be interpreted for an average adopter). Results for all remaining dependent variables are in Table D5 in Online Appendix D.

$^{+}p < 0.10$; $^{*}p < 0.05$; $^{**}p < 0.01$.

5.5. Selection on Unobservables and Robustness

If adopters of Spotify are systematically different from nonadopters in some way that is not accounted for by the observable characteristics approach in the matching procedure but that affects our dependent measures, our estimated treatment effects may be biased. As a check against this, we replicate the entire analysis using only adopters. Because of the variation of adoption time, we can use late adopters as a control for early adopters. In particular, consider an early adopter who adopts streaming at t and a late adopter who adopts streaming at $t+T$. Next, we construct a two-way fixed effects DiD estimator from comparing the difference in their behavior in periods $[0, \ldots, t-1]$ with the difference in their behavior at $[t, \ldots, t+T-1]$ (Manchanda et al. 2015).

Using this procedure, Table 8, columns (3) and (4), reports the short-run and medium-run effects of adoption on the log number of unique artists as 0.47 and 0.26, and on the log number of unique songs as 0.37 and 0.24, respectively (we cannot estimate long-run effects, since the time horizon between early and late adoption is shorter). We can compare this to the short-run and medium-run estimates, which we repeat in columns (1) and (2) of Table 8, which are 0.48 and 0.27 for artists and 0.40 and 0.26 for songs, respectively. Online Appendix D shows this holds overwhelmingly for the other outcome measures. We conclude that our main results that are based on a comparison across adopters and nonadopters are similar to results from a within-adopter analysis. This provides further support that our matching procedure is free from selection on unobservables.

We also report on a comprehensive set of additional robustness checks. Specifically, in addition to (1) controlling for unobservables using variation in adoption timing, we replicated our analysis (2) for different definitions of long-run effects (e.g., >36 weeks instead of >24), (3) excluding periods after which Taylor Swift removed content, (4) excluding the last weeks of our data before Spotify improved its recommendations and Apple Music was launched, (5) dropping countries in which Spotify was launched recently,[18] (6) focusing on consumption excluding Spotify, (7) using different functional forms (dependent variables in levels instead of logs, modeling shares; Papke and Wooldridge 1996), (8) estimating a single treatment effect instead of short-, medium-, and long-run effects, (9) using various alternative metrics for the dependent variables (e.g., top 2 and top 10 concentration, instead of Herfindahl), and (10) using all 507 adopters and 1,471 nonadopters of Spotify (i.e., without using matching). Online Appendix D provides a full overview of the

Table 8. Robustness Check with Early vs. Late Adopters: Adoption Effects on Breadth of Variety

	Main specification		Early adopters as control for late adopters	
	(1) Log number of unique artists	(2) Log number of unique songs	(3) Log number of unique artists	(4) Log number of unique songs
Short run (0–1)	0.48***	0.40***	0.47***	0.37***
	(0.040)	(0.041)	(0.057)	(0.058)
Medium run (2–24)	0.27***	0.26***	0.26***	0.24***
	(0.039)	(0.040)	(0.058)	(0.059)
R-squared	0.56	0.55	0.63	0.63
F	50.1	65.9	41.4	52.4
p-value	0.000	0.000	0.000	0.000
Users	896	896	447	447
Observations	46,688	46,688	8,445	8,445

Notes. The table shows a regression with robust standard errors in parentheses. Estimates for the main specification are calculated on a matched sample of 448 adopters and 448 nonadopters observed over 62 weeks starting May 29, 2014. User- and week-specific fixed effects are used, and the unit of analysis is the user-week. Estimates for the alternative specification are based on adopters only, with late adopters (median split; 23 weeks into the main sample) acting as a control for those who adopted earlier. The dependent variable is the log number of songs heard by a panelist in a week (play count). The independent variables are indicators for a user's time since adoption of Spotify, defined as short run (within weeks 0 and 1), medium run (within weeks 2 and 24), and long run (weeks 25 and after). Long-run effects for the alternative specification are not estimable and dropped from the table because the time between early and late adoption is shorter than 25 weeks. Complete results and robustness checks are described in Online Appendix D.

***$p < 0.001$.

results. Small numerical differences notwithstanding, all of our reported results are substantively robust.

6. Implications

We document long-run effects of streaming on music consumption: a half year after users adopt Spotify, consumption, measured in weekly play counts, is still up by 49%. By setting the price of additional variety to zero, Spotify alleviates a deadweight loss problem for varieties where valuation is positive but below the price of ownership. We also provide evidence that Spotify increases consumer welfare by reducing search frictions (e.g., enhancing discovery) and helping users discover new high-value content.

Our results point to a more fragmented market, potentially more amenable to smaller artists and labels if we rely on the representativeness of our sample. We find that Spotify adopters listen to fewer superstars and expand their attention to a wider set of artists, which could potentially increase demand for complementary goods, like live performances (Mortimer et al. 2012). The other side of the discovery coin is a drop in the staying power of songs and artists in the consumption sets of consumers. Thus, while it is easier to enter the consumption set, it is harder to stay there. Artists may thus need to exert more effort than before to stay top of mind.

7. Conclusion

The emergence of streaming services has provoked a wide-ranging debate about the benefits and drawbacks of ownership versus streaming. Constructing a unique panel data set of music consumption on streaming- and ownership-based platforms, we demonstrate the short-, medium-, and long-run effects of adoption of online streaming on quantity, variety in consumption, and new music discovery. We find that streaming increases total consumption, leads to more variety, and facilitates discovery of more highly valued music.

While our results on fragmentation in consumption pertain to the music industry, we expect similar effects to hold in other related industries like movies, TV shows, and books.

Our analysis on the effects of adopting Spotify on music consumption does not allow us to address, in satisfactory detail, the underlying mechanism that leads to the consumption changes. We postulate that they are strongly impacted by the price reduction in additional variety from adopting streaming. However, we believe that platform-specific features (e.g., personalized recommendations) may also be important. Finally, we believe that examining the role of playlists and sharing of consumption capital on music consumption are fruitful areas for future research.

Acknowledgments

The authors thank Wes Hartmann and Puneet Manchanda for comments on an earlier draft. The authors are also grateful for comments from Martin Husovec and Jeremy Silver and from seminar participants at Amsterdam Business School, INSEAD, London Business School, Tilburg University, the University of British Columbia, the University of Cologne, the Catholic University of Leuven and Vlerick Business School, the University of Michigan, the University of Oxford, Yale University, and Eurosonic Noorderslag; participants at the 2016 Marketing Science and Marketing Dynamics Conferences; and executives at Buma/Stemra, Sony Music, and Spotify.

Endnotes

[1] More broadly, we also contribute to the empirical literature on how new distribution and consumption technologies affect consumer behavior. Much of this literature is concerned with how consumption patterns change as consumers migrate online (e.g., Zentner et al. 2013), taking into account supply-side factors (e.g., greater variety online than offline) as well as demand-side factors (e.g., information on popularity). We highlight the role of another demand-side factor, the price of variety at the margin that has an impact on consumption.

[2] Interactive providers give consumers free choice in which songs they want to consume. Noninteractive providers like Pandora, instead, are similar to radio broadcasting and offer a preselected set of songs only (Aguiar 2017).

[3] For a comparison of streaming music providers, see *Wikipedia*, s.v. "Comparison of on-demand music streaming services," https://en.wikipedia.org/w/index.php?title=Comparison_of_on-demand_streaming_music_services&oldid=761564933.

[4] The service's technology monitors song-level consumption on all platforms a user has activated (on average, 4.14 platforms per user in our data). Music consumed offline is also monitored and submitted whenever a connection becomes available. Traditional FM radio is not included.

[5] 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 Z_i, defined following Equation (1).

[6] We retrieved genre data from Echonest.com (the industry's leading music intelligence provider, using the service's API) and Musicbrainz.org (an open-source music encyclopedia, using a virtual replication of their SQL server). This information was available for 53% of all artists.

[7] The "other" category consists of the following players (listed if their market share is >1%): the Service Plugin (7.99%), Android Service (4.33%), Simple Service Plugin (2.28%), Service plugin for web (1.33%), plugin for iOS (1.23%), and Clementine (1.08%). The remaining 211 niche players account for 8.4% of total play counts. YouTube is part of our data, but largely tracked in combination with other web-based players. Hence, we cannot reliably distinguish YouTube from other services such as SoundCloud.

[8] We ignore other streaming platforms in our sample because of their low market shares: Grooveshark (0.54%), Rdio (0.42%), Google Play (0.17%), simfy (0.07%), WiMP (0.03%), Deezer (0.01%), Tidal (0.01%), and 8 tracks (0.01%).

[9] See Figure C1 in Online Appendix C for the distribution of Spotify adoption times.

[10] Further insight is obtained from Alexa's audience demographics for Spotify.com (available at http://www.alexa.com/siteinfo/spotify.com; accessed January 25, 2017). Assuming an equal distribution of males and females on the Internet, Alexa estimates visitors to Spotify.com being about 60% male and 40% female.

[11] The average weekly number of songs that Spotify adopters listen to, conditional on listening in a given week, is 105 songs. Assuming an average song length of four minutes, we can compute the daily usage of Spotify to be 105 songs × 4 minutes = 420 minutes per week = 60 minutes per day.

[12] As explained in Section 3.2, we have excluded skipped songs from our data.

[13] As a robustness check we also perform all analyses without any matching (see Online Appendix D). We find similar results.

[14] We do not directly observe whether listeners use the free or the premium version of Spotify. However, the free version features 30-second advertisements and the premium version does not. We know the length for each song and the time elapsed between starting two adjacent songs. This allows us to compute a metric on the frequency of 25- to 35-second (putatively, advertising) gaps relative to the number of adjacent song pairs. We use this metric as a proxy for which consumers have the free version with advertisements. Assuming consumers listen to 15 songs an hour, the median user is exposed to (0.0374 gaps per adjacent song pair × 15 songs) = 0.561 ads an hour.

[15] Our data allow us to estimate these long-term effects reliably. For example, we observe 296 adopters for more than 24 weeks, 223 for more than 36 weeks, 131 for more than 48 weeks, and 85 for more than 52 weeks.

[16] Spotify and iTunes/WWF may be compliments even if the estimated iTunes/WWF consumption effects are negative. For example, users may make fewer (but better) purchase decisions on iTunes after sampling content on Spotify. Although we acknowledge that this may happen in some rare cases, our data do not provide support for this interpretation. First, the large majority of tracks discovered on Spotify are consumed on Spotify itself (98.35% share of plays); only 0.48% of all plays for these discoveries occur on iTunes, 0.43% on WWF, and 0.74% on other services. Second, consumers do not appear to make better consumption choices outside of Spotify. In fact, using repeat consumption not on Spotify as a proxy for consumer-preference fit, we find that it decreases rather than increases (see new content metrics in Table D5 in Online Appendix D). Hence, there is little evidence in our data that suggests that consumers purchase content on iTunes after sampling it on Spotify.

[17] To what extent does the content removal of a very popular artist (Taylor Swift) during our observation period affect our results? Confining our analysis to the period *before* her content removal (see Online Appendix D, Table D4), we find that the effects tend to be stronger. Hence, a decrease in superstar consumption is not likely driven by Taylor Swift. We thank a reviewer for making this suggestion.

[18] Spotify was recently launched in Brazil and Canada. The (early) adopters from these countries arguably have a higher interest in variety compared to adopters from other countries captured relatively late in the diffusion of Spotify. Hence, findings for Brazil and Canada could potentially be informative about the degree to which our main sample (with the majority of users coming from outside of Brazil and Canada) is skewed toward users with a lower need for variety. We in fact find differences in the adoption effects, which are generally more pronounced in Canada and Brazil compared to the remaining countries (see Online Appendix D). However, there are several explanations for this difference. One is that there is heterogeneity in taste for variety across countries. Another is that all countries have the same distribution of taste for variety, but that different countries are at different stages of diffusion. Since we do not have data to identify the difference between these scenarios, we caution against interpreting these effects as purely driven by users' interest in variety.

References

Adler M (1985) Stardom and talent. *Amer. Econom. Rev.* 75(1):208–212.

Aguiar L (2017) Let the music play? Free streaming and its effects on digital music consumption. *Inform. Econom. Policy* 41:1–14.

Aguiar L, Waldfogel J (2015) Streaming reaches flood stage: Does Spotify stimulate or depress music sales? Working paper, National Bureau of Economic Research, Cambridge, MA.

Alexander P (1997) Product variety and market structure: A new measure and a simple test. *J. Econom. Behav. Organ.* 32(2): 207–214.

Bertrand M, Duflo E, Mullainathan S (2004) How much should we trust differences-in-differences estimates? *Quart. J. Econom.* 119(1):249–275.

Bronnenberg B (2015) The provision of convenience and variety by the market. *RAND J. Econom.* 46(3):480–498.

Bronnenberg B, Dubé JP, Mela C (2010) Do digital video recorders influence sales? *J. Marketing Res.* 47(6):998–1010.

Cameron S, Collins A (1997) Transaction costs and partnerships: The case of rock bands. *J. Econom. Behav. Organ.* 32(2):171–183.

Chung K, Cox A (1994) A stochastic model of superstardom: An application of the Yule distribution. *Rev. Econom. Statist.* 76(4): 771–775.

Chung T, Rust R, Wedel M (2009) My mobile music: An adaptive personalization system for digital audio players. *Marketing Sci.* 28(1):52–68.

Crain W, Tollison R (2002) Consumer choice and the popular music industry: A test of the superstar theory. *Empirica* 29:1–9.

Elberse A (2008) Should you invest in the long tail? *Harvard Bus. Rev.* 86(7–8):88–96.

Friedlander JP (2016) News and notes on 2015 RIAA shipment and revenue statistics. https://www.riaa.com/wp-content/uploads/2016/03/RIAA-2015-Year-End-shipments-memo.pdf.

Gensler S, Leeflang P, Skiera B (2012) Impact of online channel use on customer revenues and costs to serve: Considering product portfolios and self-selection. *Internat. J. Res. Marketing* 29(2): 192–201.

Hamlen W (1991) Superstardom in popular music: Empirical evidence. *Rev. Econom. Statist.* 73(4):729–733.

Hoch SJ, Bradlow ET, Wansink B (1999) The variety of an assortment. *Marketing Sci.* 18(4):527–546.

Holbrook M, Hirschman E (1982) The experiential aspects of consumption: Consumer fantasies, feelings, and fun. *J. Consumer Res.* 9(2):132–140.

Kim J, Allenby GM, Rossi PE (2002) Modeling consumer demand for variety. *Marketing Sci.* 21(3):229–250.

Kissel C (2015) Spotify listeners discover roughly 27 new artists a month. *Diffuser* (July 23), http://diffuser.fm/spotify-listeners-discover-roughly-27-new-artists-a-month/.

Lacher KT, Mizerski R (1994) An exploratory study of the response and relationships involved in the evaluation of, and in the intention to purchase new rock music. *J. Consumer Res.* 21(2): 366–380.

Lopes P (1992) Innovation and diversity in the popular music industry, 1969 to 1990. *Amer. Sociol. Rev.* 57(1):56–71.

Manchanda P, Packard G, Pattabhiramaiah A (2015) Social dollars: The economic impact of customer participation in a firm-sponsored online customer community. *Marketing Sci.* 34(3): 367–387.

Mitroff S, Blanco X (2015) Apple Music vs. Spotify: What's the difference? *CNET* (July 2), https://www.cnet.com/news/apple-music-vs-spotify-whats-the-difference/.

Mortimer J, Nosko C, Sorensen A (2012) Supply responses to digital distribution: Recorded music and live performances. *Inform. Econom. Policy* 24(1):3–14.

Nielsen (2014) Nielsen Music U.S. Report. http://www.nielsen.com/content/dam/nielsenglobal/kr/docs/global-report/2014/2014%20Nielsen%20Music%20US%20Report.pdf.

Oestreicher-Singer G, Zalmanson L (2013) Content or community? A digital business strategy for content providers in the social age. *MIS Quart.* 37(2):591–616.

Papke LE, Wooldridge JM (1996) Econometric methods for fractional response variables with an application to 401 (K) plan participation rates. *J. Appl. Econometrics* 11(6):619–632.

Ratner RK, Kahn BE, Kahneman D (1999) Choosing less-preferred experiences for the sake of variety. *J. Consumer Res.* 26(1):1–15.

Rob R, Waldfogel J (2006) Piracy on the high C's: Music downloading, sales displacement, and social welfare in a sample of college students. *J. Law Econom.* 49(1):29–62.

Rosen S (1981) The economics of superstars. *Amer. Econom. Rev.* 71(5):845–858.

Rosenbaum P, Rubin DB (1983) The central role of the propensity score in observational studies for causal effects. *Biometrika* 70(1):41–55.

Spotify (2016) Streaming state of mind. https://spotifyforbrands/.com/us/insight/streaming-state-of-mind/.

Spotify (2017a) Fast facts. https://press.spotify.com/us/about/.

Spotify (2017b) Spotify for brands. https://spotifyforbrands.com/us/.

van Herpen E, Pieters R (2002) The variety of an assortment: An extension to the attribute-based approach. *Marketing Sci.* 21(3): 331–341.

Waldfogel J (2012) Music piracy and its effects on demand, supply, and welfare. *Innovation Policy Econom.* 12(1):91–110.

Wlömert N, Papies D (2016) On-demand streaming services and music industry revenues—Insights from Spotify's market entry. *Internat. J. Res. Marketing* 33(2):314–327.

Zentner A (2006) Measuring the effect of online piracy on music sales. *J. Law Econom.* 49(1):63–90.

Zentner A, Smith M, Kaya C (2013) How video rental patterns change as consumers move online. *Management Sci.* 59(11):2622–2634.