What Makes Popular Culture Popular?: Product Features and Optimal Differentiation in Music

Noah Askin and Michael Mauskapf

INSEAD, Fontainebleau, France
noah.askin@insead.edu

Columbia Business School, New York, NY
mm5123@columbia.edu

Forthcoming at *American Sociological Review*

ABSTRACT

In this paper, we propose a new explanation for why certain cultural products outperform their peers to achieve widespread success. We argue that products' positioning within feature space significantly predicts their popular success. Using tools from computer science, we construct a novel data set that allows us to test how the musical features of nearly 27,000 songs from *Billboard's* Hot 100 charts structure the consumption of popular music. We find that, in addition to artist familiarity, genre affiliation, and institutional support, a song's perceived proximity to its peers influences its position on the charts. Contrary to the claim that all popular music sounds the same, we find that songs sounding too much alike—those that are highly typical—are less likely to succeed, while those exhibiting some degree of optimal differentiation are more likely to rise to the top of the charts. These findings offer a new contingent perspective on popular culture by specifying how content organizes competition and consumption behavior in cultural markets.

Keywords: consumption, music, optimal differentiation, popular culture, product features, typicality

For their valuable ideas, comments, and support, we are grateful to Matt Bothner, Frédéric Godart, Alejandro Mantecón Guillén, Brayden King, Jenn Lena, John Levi Martin, John Mohr, Klaus Weber, Erich Ludwig and the rest of the Echo Nest team, and members of the University of Chicago's Knowledge Lab and Northwestern University's Institute on Complex Systems (NICO). We also thank the *ASR* editorial team and three anonymous reviewers for their thoughtful and thought-provoking feedback, which made the paper better.

INTRODUCTION

What makes popular culture popular? Scholars across the humanities and social sciences have spilled considerable ink trying to answer this question, but our understanding of why certain cultural products succeed over others remains incomplete. Although popular culture tends to reflect, or is intentionally aimed toward, the tastes of the general public, there exists wide variation in the relative popularity of these products (Rosen 1981; Storey 2006). Extant research in sociology and related disciplines suggests that audiences seek and utilize diverse information that might signal the quality and value of new products (Keuschnigg 2015), including the characteristics and networks of cultural producers (Peterson 1997; Uzzi and Spiro 2005; Yogev 2009), audience preferences and social influence dynamics (Lizardo 2006; Mark 1998; Salganik, Dodds, and Watts 2006), elements in the external environment (Peterson 1990), and various institutional forces (Hirsch 1972).

Each of these signals plays an important role in determining which products audiences select, evaluate, and recommend to others. Nevertheless, while these choices and the preferences they represent vary widely over time and across individuals, research suggests that the inherent quality of cultural products also affects how audiences classify and evaluate them (Goldberg, Hannan, and Kovács 2015; Jones et al. 2012; Lena 2006; Rubio 2012; Salganik et al. 2006). Certain product features may independently signal quality and attract audience attention (e.g., Hamlen 1991), but we believe that these features matter most *in toto*, both by creating a multi-dimensional representation of products and by positioning those products across the plane of possible feature combinations.[1] Rather than existing in a vacuum, cultural products are perceived in relation to one another, and these relationships shape how consumers organize and discern the art worlds around them (Becker 1982).

One way to think about how product position shapes performance outcomes is through the lens of categories research, which highlights how social classification systems organize consumers' expectations and preferences (Hsu 2006; Zuckerman 1999) and help them draw connections between products. We agree that categories play a significant role in structuring taste and consumption behavior (Bourdieu 1993), but much of the work in this area makes the implicit assumption that category *labels* remain tightly coupled with a set of underlying *features*. Recent research notes, however, that these features may not cluster or align with prevailing classification schemes (Anderson 1991; Kovacs and Hannan 2015; Pontikes and Hannan 2014).[2] Category labels (e.g., "country" in the case of musical genres) work well when navigating stable product markets with clearly defined category boundaries, but they do not always reflect how audiences actually make sense of the world in which they are embedded, especially in contexts where products are complex and tastes are idiosyncratic and dynamic (Lena 2015). In these domains, extant categories may not provide adequate or accurate information to consumers, who must instead rely on the underlying features of products to draw comparisons and make selection decisions.

We build on these insights to propose a new explanation for why certain cultural products outperform their competitors to achieve widespread success. In the context of popular music, we argue that audiences use musical signals to draw latent associations between songs. These associations, which are conceived independently from traditional categories, help to organize the choice set from which audiences select and evaluate products, positioning certain songs more advantageously than others. We hypothesize that hit songs are able to successfully manage a similarity-differentiation tradeoff, simultaneously invoking conventional feature combinations associated with previous hits while inciting some degree of novelty that distinguishes them from their peers. This prediction speaks to the competitive benefits of optimal differentiation, a finding that reoccurs across multiple studies and areas in sociology and beyond (Lounsbury and Glynn 2001; Uzzi et al. 2013; Zuckerman 2016).

To test this prediction and better understand the relationship between product features and success in music, we construct a novel dataset consisting of nearly 27,000 songs that appear on the *Billboard* Hot 100 charts between 1958 and 2016. The data include algorithmically-derived features that describe a song's sonic quality. Sonic features range from relatively objective musical characteristics, such as "key," "mode," and "tempo," to more perceptual features that quantify a song's "acousticness," "energy," and "danceability," among others. After demonstrating the baseline significance of individual features in predicting a song's peak position and longevity on the charts, we use these features to construct a measure of sonic similarity or typicality and test its effect on chart performance. While popular opinion suggests that songs are most likely to succeed when they adhere to a conventional and reproducible template (Dhanaraj and Logan 2005; Thompson 2014), we find that the most successful songs in our dataset are optimally differentiated from their peers. Our results provide strong evidence that, net of other factors such as artist familiarity and genre affiliation, product features matter, particularly in the way they structure songs' relationships to each other. These findings, and the data and methods we use to produce them, make important contributions to economic sociology and the sociology of culture by offering a new contingent perspective on popular culture. Using new, micro-level feature data to specify how cultural content organizes the way in which audiences distinguish products compels us to rethink some of the basic mechanisms behind consumption and taste formation.

CULTURAL PREFERENCES AND THE SIMILARITY-DIFFERENTIATION TRADEOFF

Predicting how well a new product will fare in the marketplace for audience attention presents a difficult, if not impossible, challenge, due to the countless variables and contingencies that may influence performance outcomes. This challenge is particularly pronounced in the realm of the cultural or "creative" industries (Caves 2000; Hadida 2015), which tend to generate products and experiences whose evaluation involves significant subjectivity (Krueger 2005). Even after a cultural product—a painting, film, or song—has been anointed a "success," it can be difficult to explain *ex post* why certain products enjoy more success than others (Bielby and Bielby 1994; Lieberson 2000). The relative popularity of a cultural product is usually ascribed to prevailing tastes, which are largely considered a function of individuals' idiosyncratic preferences, past experiences, and exposure patterns, as well as the prevailing opinions of others. Moreover, different types of performance outcomes (e.g., mass appeal vs. critical acclaim) beget different varieties of explanation, and require audiences to consider distinct dimensions of evaluation that are often context specific.[3] Our ability to explain what constitutes a hit versus a flop remains limited.

Scholars interested in this question have traditionally taken one of several approaches to explain the determinants of cultural preferences and product performance. The first set of explanations focuses on the characteristics of cultural producers, including their reputation (Bourdieu 1993), past performance outcomes (Peterson 1997), and the structure of their professional networks (Godart, Shipilov, and Claes 2014; Yogev 2009). Indeed, just as cultural products are perceived by audiences in relation to one another, they are also created by producers who form collaborative relationships and draw inspiration from each other's work. In the context of Broadway musicals, Uzzi and Spiro (2005) find that when collaborations between artists and producers display small world properties, their cultural productions are more likely to achieve critical and commercial success. Phillips (2011, 2013) finds that the artists who are most likely to re-record and release jazz standards come, surprisingly, from structurally disconnected cities. Research on sampling in rap music (Lena and Pachucki 2013), innovations in video game production (de Vaan, Stark, and Vedres 2015), and the creative success of inventors (Fleming, Mingo, and Chen 2007) provides ample evidence that certain types of producer networks are more likely to generate new and successful products through the recombination of diverse ideas. Thus, the interconnectedness of producers and of the production process more generally plays an important role in shaping product performance and consumer taste.

It is worth noting here that channels of influence between networks and taste run in both directions (Lizardo 2006). Just as social networks can alter cultural outcomes, so too can those networks be altered by prevailing tastes and practices, recasting culture and social structure as mutually constitutive (Pachucki and Breiger 2010; Vaisey and Lizardo 2010). This view—one that highlights culture's role in determining social reality—is supported by the "strong program" in cultural sociology (e.g., Alexander and Smith 2002) and related work on the materiality of culture (Rubio 2012). Rather than passive symbolic structures, culture is endowed with real properties that can influence actors' preferences, behaviors, and affiliation patterns.

The second set of variables used to explain the success of cultural products pertains to audience or demand-side characteristics. Variables of this sort include individual and collective trends in demand, as well as other related consumer dynamics, such as homophily (Mark 1998) and endogenous diffusion patterns (Rossman 2012). These explanations speak to the significant role of social influence, which is often responsible for wide variances in product adoption and taste formation (DellaPosta, Shi, and Macy 2015). In a series of online experiments, Salganik and colleagues investigated how product quality and social influence affect success in an artificial music market (Salganik et al. 2006; Salganik and Watts 2008; Salganik and Watts 2009). Despite the outsized role of social influence, they found compelling evidence that the likelihood of a song being downloaded by participants is determined in part by its inherent quality—but the exact nature of such "quality" remains a mystery.

The categories literature provides a third class of explanations for the variable success of cultural products (Hsu 2006; Jones et al. 2012). Product categories and the labels attached to them reflect largely agreed-upon conventions that audiences attribute to certain groups of products. In this sense, "products are cultural objects imbued with meaning based on shared understandings, and are themselves symbols or representations of those meanings" (Fligstein and Dauter 2007). Much of the research on social classification explores the role of categories in organizing product markets and consumer choice. This process is particularly salient in cultural markets (Caves 2000; DiMaggio 1987), where classification systems provide the context through which producers and consumers structure their tastes, preferences, and identities (Bourdieu 1993; Peterson 1997), and determine how they search and evaluate the art worlds around them (Becker 1982). Indeed, the emergence and institutionalization of genre categories features prominently in explanations of market competition across a number of cultural domains, including film (Hsu, 2006), painting (Wijnberg and Gemser 2000), literature (Frow 1995, 2006), and music (Frith 1996; Holt 2007; Lena and Peterson 2008; Negus 1992).

Category researchers have made considerable contributions to our understanding of when and why certain kinds of organizations or products succeed (Hsu, Negro, and Perretti 2012; Zuckerman 1999), but this work has several important limitations. Although categories play an important role in shaping how audiences search, select, and evaluate products, they often provide a relatively coarse and static picture of “the market,” assuming a nested hierarchical structure that is more or less agreed-upon by market actors. We know, however, that categories and their boundaries are dynamic and eminently contested, signifying different meanings to different communities (Lena 2012; Sonnett 2004). Moreover, most research in this area highlights the social-symbolic labels attached to categories, ignoring the high-dimensional features of the products that occupy them. While labels constitute socially constructed and symbolically ascribed descriptors for a given category, features provide considerably more fine-grained information about a focal product’s underlying composition and position in “conceptual space” (Kovács and Hannan 2015). Recent research indicates that individuals classify products and other entities across a number of different dimensions, including shared cultural frames or world views (Goldberg 2011), overlapping cognitive interpretations (De Vaan, Stark, and Vedres 2015), and interpersonal connections between producers or consumers (Lena 2015). The classification structures that emerge from these processes may or may not align with explicit categorical prescriptions such as musicological genre, suggesting an alternative dictum by which audiences position and compare similar producers and their products in the marketplace.

Product Features and Audience Associations

Category labels are usually coupled with a set of underlying features or attributes, but the degree of coupling between features and labels is highly variable (Anderson, 1991; Pontikes and Hannan 2014). For example, Bob Dylan’s version of “Like a Rolling Stone” might be tagged with labels like “Folk,” “Americana,” or even “Rock-N-Roll,” but it also exhibits countless features, including its duration (6:09), key (C Major), instrumentation (vocals, guitar, bass, electric organ, harmonica, tambourine), and thematic message (love, resentment). From our perspective, these features—the inherent, high-dimensional attributes that constitute the “DNA” of individual products—are culturally determined, grounding products in material reality and granting them structural autonomy (Alexander and Smith 2002). Recent research suggests that the features of cultural products also shape classification processes and performance outcomes (Jones et al. 2012; Lena 2006; Rossman and Schilke 2014). Like category labels, features can be used to position products that seem more or less similar to each other (see Cerulo 1988), shaping consumers' perceptions and sensemaking in distinct ways (Tversky 1977). Furthermore, empirical evidence from popular music suggests that certain features (e.g., instrumentation) shape listening preferences and play an important role in determining why some products succeed and others fail (Nunes and Ordanini 2014).

Our reading of these literatures suggests that there is a gap in the way we conceptualize features and their role in positioning products for success. Rather than influencing consumption independently, we believe that features cohere in particular combinations to generate holistic, *gestalt* representations of products. Recent work at the vanguard of network neuroscience is beginning to explore how individuals cognitively structure and make sense of these representations (Brashears and Quintane 2015; Zerubavel et al. 2015), but we still know little about how this process unfolds.[4] In the context of consumption, we argue that consumers position products across some multi-dimensional feature space, whereby certain objects are perceived to be more (or less) similar depending on the features they do (or do not) share. These latent associations represent the world of products in which consumers are embedded, and exhibit a social life all their own (Carroll, Khessina, and McKendrick 2010; Douglas and Isherwood 1996).[5] They also organize the relevant comparison set from which consumers select and evaluate products.

This argument is distinctive in several important ways. First, we highlight the consequentiality of the implicit relationships formed between products, rather than producers, consumers, or category labels. Audience evaluations of products are shaped not only by the characteristics of producers and consumers, or social influence pressures, but also by a product's position within a broader ecosystem of cultural production and consumption. The intuition behind this argument is relatively straightforward: while the choices consumers make are shaped by their individual preferences, relationships, and various other factors, they are also influenced by the feature-based similarity space within which products are embedded (Kovács and Hannan 2015). Put another way, a consumer's direct and indirect exposure to some set of related products plays a critical role in shaping his or her future selection decisions and preferences.

Second, we argue that the structure and effect of these feature-based associations are conceptually and analytically distinct from those usually attributed to traditional categories. Research on category emergence suggests that labels and features operate across separate planes, which may or may not align with one other (Pontikes and Hannan 2014). We already know that consumers refer to established categories to make sense of the products they encounter (e.g., Zuckerman 1999), but recent work at the intersection of culture, cognition, and strategy identifies the distinctive role of "product concepts," which form loose relational structures that shape consumer cognition beyond purely categorical classification (Kahl 2015). These insights reinforce our interest in feature-based associations, and suggest that consumers in certain contexts are likely to use an amalgamation of features rather than (or in addition to) labels to position, select, and evaluate products. In the analysis that follows, we account for both of these dimensions to explain why certain songs attract audience attention and outperform their competition in the market for popular music.

The Similarity-Differentiation Tradeoff

We have already reviewed a number of plausible explanations for the variable success of cultural products, including producer reputation and category membership, but the study of product features and the associations they generate provides a new set of mechanisms to explain why certain products achieve popularity while others do not. One common way to examine the effects of product positioning on market performance is to measure crowding and differentiation dynamics (e.g., Bothner, Kang, and Stuart 2007). This strategy has been particularly useful in the organizational ecology literature (Podolny, Stuart, and Hannan 1996; Barroso et al. 2014), where the presence of too many competitors can saturate a consumer or product space (e.g., niche), making it increasingly difficult for new entrants to survive. Research across a number of empirical contexts suggests that the ability to differentiate oneself and develop a distinctive identity can help products, organizations, and other entities compete within or across niches (Hannan and Freeman 1977; Hsu and Hannan 2005; Swaminathan and Delacroix 1991). Alternatively, some work in cognitive and social psychology argues that conformity is the recipe for success. For example, research on liking (Zajonc 1968) suggests that the more people are exposed to a stimulus, the more they enjoy it, regardless of whether or not they recognize having been previously exposed. In music, this means that the more a song sounds like something the listener has heard before, the more likely they are to evaluate it positively and listen to it again (see Huron 2013). This argument lies at the heart of "hit song science," which claims that, with enough marketing support, artists can produce a hit song simply by imitating past successes (Dhanaraj and Logan 2005; Thompson 2014).

Rather than test these competing predictions individually, we hypothesize that the pressures toward conformity and differentiation act in concert. Products must differentiate themselves from the competition to avoid crowding, but not so much as to make themselves unrelatable (Kaufman 2004). Research on consumer behavior suggests that audiences simultaneously pursue these competing goals as well, conforming on certain identity-signaling attributes (such as a product's brand or category) while distinguishing themselves on other product features (such as color or instrumentation; see Chan, Berger, and Van Boven 2012). This tension between differentiation and conformity is central to our understanding of social identities (Brewer 1991), category spanning (Zuckerman 1999; Hsu 2006), storytelling (Lounsbury and Glynn 2001), consumer products (Lancaster 1975), and taste (Lieberson 2000). Taken together, this work signals a common trope across the social sciences: the path to success requires some degree of both conventionality and novelty (Uzzi et al. 2013).

In the context of popular music, we expect that songs able to strike a balance between "being recognizable" and "being different"—those that best manage the similarity-differentiation tradeoff—will attract more audience attention and experience more success. Stated more formally, we predict an inverted U-shaped relationship between a song's relative typicality and performance on the *Billboard* Hot 100 charts. Our analysis highlights the opposing pressures of crowding and differentiation by constructing a summary measure of song typicality, which accounts for how features position a song relative to its musical neighbors. Controlling for a host of other factors, including an artist's previous success and genre affiliation, we expect that songs exhibiting *optimal differentiation* across the feature space are more likely to achieve widespread popularity, while those that are deemed too similar to—or dissimilar from—their peers will struggle to reach the top of the charts (cf. Zuckerman 2016).

DATA & METHODOLOGY

Studying the relative typicality of products can shed light on how audience preferences are shaped across a number of empirical contexts, but we believe music represents an ideal setting in which to test these dynamics, due in part to its reliance on an internally consistent grammar. While songs can be quite different from one another, they follow the same set of basic "rules" based on melody, harmony, and rhythm; listeners' tastes, on the other hand, do not have such concrete bounds. Although Salganik and colleagues (2006) showed that consumer choice in an artificial music market is driven both by social influence *and* a song's inherent quality, their measure of "quality" is simply audience preference in the absence of experimenter manipulation. Measuring quality "objectively" requires a comprehensive technical understanding of music's form and features. Due to the specialized skills needed to identify, categorize, and evaluate such features reliably, work that meets these demands is limited. The research that has been conducted employs musicological techniques to construct systems of comparable musical codes that may be more or less present in a particular musical work (Cerulo 1988; La Rue 2001; Nunes and Ordanini 2014). Yet even if social scientists learned these techniques, or collaborated more often with musicologists, it would be very difficult to apply and automate such complex codes at scale.

Fortunately, these difficulties have been partially attenuated by the application of digital data sources and new computational methods to the study of culture. Developed first by computer scientists and then adopted by mainstream social science, these technologies have begun to filter into the toolkits of cultural sociologists (Bail 2014), who have traditionally been criticized for being "methodologically impoverished" (DiMaggio, Nag, and Blei 2013). Most relevant for our purposes are advances in music information retrieval (MIR) and machine learning (e.g., Friberg et al. 2014; Serrà et al. 2012), fields that have developed new methods to reduce the high dimensionality of musical compositions to a set of discrete features, much like topic modeling has done for the study of large texts (Blei, Ng, and Jordan 2003). These developments have generated new research possibilities that were previously considered impractical. Using a novel dataset that includes discrete representations of musical features in the form of sonic features (a song's "acoustic footprint"), we investigate how popular success is contingent in part on a song's relative position within feature space.

Our primary data come from the weekly *Billboard* Hot 100 charts, which we have reconstructed from their inception on August 4, 1958 through March 26, 2016. The Hot 100 charts are published by *Billboard Magazine*, but the data we use for our analysis come from an online repository known as "The Whitburn Project." Joel Whitburn collected and published anthologies of the charts (Whitburn 1986, 1991) and, beginning in 1998, a dedicated fan base started to collect, digitize, and add to the information contained in those guides. This augmented existing chart data, providing additional details about the songs and albums on the charts, including metadata and week-to-week rankings for more than 26,800 songs spanning almost 60 years. A descriptive comparison of these songs with others that did not appear on the Hot 100 charts suggests that, while the observations included in our analysis constitute a slightly more homogenous or "typical" sample than is represented in music broadly, *the distribution of song typicality across these samples is nearly identical*, making the charts an appropriate proxy for studying consumer evaluation and product performance in the field of popular music (see Appendix A for a more formal comparison).[6] Furthermore, although the algorithm used to create the charts has evolved over the years—something we account for in our analysis—they remain the industry standard.[7] As such, they have been used extensively in social science research on popular music (Alexander 1996; Anand and Peterson 2000; Bradlow and Fader 2001; Dowd

2004; Lena 2006; Lena and Pachucki 2013; Peterson and Berger 1975), and are noted for their reliability as indicators of popular taste (e.g., Eastman and Pettijohn II 2014).

Dependent Variables. The weekly *Billboard* charts provide us with a real-world performance outcome that reflects the general popularity of a song and can be tracked and compared over time. Unlike movie box-office results or television show ratings, music's content owners closely guard sales data, leaving songs' diffusion across radio stations (Rossman 2012) or their chart position as the most reliable and readily available performance outcome. In their examination of fads in baby naming, Berger and Le Mens (2009) use both peak popularity and longevity as key variables in the measurement of cultural diffusion; we adapt them here as our dependent variables, *peak position* and *weeks on charts*. Although these two outcomes are related to one another (i.e., songs that reach a higher peak chart position tend to remain on the charts longer, $R \approx .72$), we test both measures in our analysis. We also reverse code peak chart position (101 – chart position) so that positive coefficients on our independent variables indicate a positive relationship with a song's success on the charts.

To account for the competitive dynamics between songs appearing on the same chart, we also include a set of models that employ a third measure of success based on week-to-week change in chart position. We subtracted each song's (reverse-coded) position during the previous week (t) from its current position (t+1) to determine the effect of song typicality on weekly changes in chart position. While a third dependent variable complicates our analysis, we believe this approach is appropriate because it (1) better captures the dynamic nature of the charts, which can change considerably from week to week, while allowing us to include fixed effects for songs; (2) does not penalize the relatively short "shelf life" of song popularity; and (3) accounts for the fact that songs appearing near the bottom of the charts have greater opportunity for improvement when compared to those at the top.

Genre Data. The *Billboard* data require augmentation to capture more fully the multifaceted social and compositional elements of songs and artists. Although genre categories evolve and are potentially contentious (Lena and Pachucki 2013), they provide an important form of symbolic classification that organizes the listening patterns and evaluations of producers, consumers, and critics (Bourdieu 1993; Holt 2007; Lena 2012). Moreover, genres play a significant role in defining and shaping artists' identities (e.g., Peterson 1997; Phillips and Kim 2008), which in turn help to determine the listeners who seek out and are exposed to new music. Audiences consequently reinforce artist identities and genre structures (Negus 1992; Frith 1996), setting expectations for both producers and their products.

To account for the effect of traditional category labels, we collected musicological genre data from Discogs.com, an encyclopedic music site and marketplace containing extensive information on music recordings, specifically singles and albums (see Montauti and Wezel 2016). Like other music websites, particularly those with user-generated and curated data, Discogs includes multiple genre and style (or sub-genre) attributions for each release (i.e., single, album, EP or LP). Although up to three genre and six style attributions are possible, we created dummy variables for the *primary genre* affiliated with each release in our analysis (see "crossovers" below for an exception). Many songs on the Hot 100 were released as singles, allowing us to obtain fine-grained, song-level genre classification data. For those songs that were not released as singles, we use the primary genre attributed to the album on which the song appears.[8] Based on these data, our sample covers fifteen genre categories—including Pop, Rock, Blues, Electronic, Jazz, and Hip Hop.[9]

Echo Nest Sonic Feature Data. Although genre represents an important means of symbolic classification in music, our interest in more fine-grained, feature-based associations necessitated the collection of data summarizing the sonic attributes of each song. For these data we turned to the Echo Nest, an online music intelligence provider that offered access to much of their data via a suite of Application Programming Interfaces (APIs). This organization represents the current gold standard in MIR, having been purchased by music streaming leader Spotify in 2014 to power its analytics and recommendation engines. Using web crawling and audio encoding technology, the Echo Nest has collected and continuously updates information on over 30 million songs and 3 million artists. Their data contains objective and derived qualities of audio recordings, as well as qualitative information about artists based on text analyses of artist mentions in digital articles and blog posts.

We accessed the Echo Nest API to collect complete data on 94% of the songs (25,102 of 26,846 total songs) that appeared on the charts between 1958 and 2016, including several objective musical features (e.g., “tempo,” “mode,” and “key”), as well as some of the company’s own creations (e.g., “valence,” “danceability,” and “acousticness”). Songs are assigned a quantitative value for each feature, which are measured using various scales. **Table 1** briefly describes the eleven features used in our analysis. There are of course limitations associated with distilling complex cultural products into a handful of discrete features, but we believe that these features represent the best available approximation of what people hear when they listen to music. Nearly twenty years of research and advancements in MIR techniques have produced both high- and low-level audio features that provide an increasingly robust representation of how listeners’ perceive music (Friberg et al. 2014). Our conversations with leading MIR researchers support our belief that these measures provide the most systematic attempt to capture songs’ material and sensory composition at scale. Moreover, these features were created specifically for song-to-song comparisons to inform algorithmically-generated recommendations for listeners.

[Table 1 around here]

Independent Variable: Song Typicality. In an effort to provide a more nuanced explanation of how a song's relative position within feature space affects performance, we construct a dynamic measure of song typicality. For this variable, *genre-weighted typicality (yearly)*, we measure the cosine similarity between songs using the sonic features provided by the Echo Nest—normalizing each to a 0-1 scale so as to not allow any individual attribute undue influence over our similarity calculation, and then collapsing them into a single vector $\boldsymbol{V_i}$ for each song in our dataset.[10] For each song *i*, we pulled every other song that appeared on the charts during the year prior to song *i*'s debut, and calculated the cosine similarity between each song-pair's vector of features. The resulting vector $\mathbf{V}_{it}$ includes the cosine similarity between song *i* and every other song *j* from the previous 52 weeks' charts, which we consider the "boundary" of the relevant comparison set against which each song is competing.

After thoughtful consideration, we determined that simply taking the average of each song's row of similarities in $\mathbf{V}_{it}$—in essence, creating a summary typicality score for each song in our dataset—left open the possibility that two songs which "looked" similar (in terms of their constitutive features) might actually sound different, thus biasing our analysis. Furthermore, research suggests that consumers tend to be split into segments defined by the type of music that they consume. These segments or communities may or may not align with traditional "musicological" genre categories, which have their own distinct traditions and histories (cf. Lena

2012, 2015). Although omnivorous consumption behavior is on the rise (e.g., Lizardo 2014), we believe that the perceived sonic similarity between two songs will decrease if those songs are associated with different genres (e.g., a country song and a reggae song may have similar beat and chord structures, thereby “appearing” to be similar when seen as a vector of features, but perceived to “sound” quite different by listeners). Thus, we weight each song-pair's raw cosine similarity by the average similarity of those songs’ parent genres over the preceding 52 weeks.

We chose to use a genre-weighted cosine similarity measure for two reasons. First, we wanted to generate a fine-grained, feature-based measure, rather than one based purely on shared symbolic classification. Although the latter represents an important signal of how listeners identify and process music, we focus on the former because we believe it provides a more objective representation of a song’s sonic fingerprint. Moreover, cosine similarity is a common measure for clustering multi-dimensional vectors (Evans 2010). Second, we believed it was important to include all songs in a given year, rather than only songs from within a particular genre, when constructing a relevant comparison set to measure typicality. Listeners may be more likely to listen to and compare songs from within the same genre—this is why we chose to incorporate a genre weighting scheme in the first place—but we also recognize that for many listeners these genres and their boundaries are not absolute, particularly when it comes to the most mainstream music being captured on the Hot 100 charts. We therefore decided to include songs from all genres when defining the relevant comparison set for our main typicality measure.

To construct the weights for our *genre-weighted typicality (yearly)* measure, we calculated yearly within-genre averages for each sonic feature, and then again used a cosine similarity algorithm to measure the average proximity of each pair of genres in feature space. The resulting similarities were then applied to the raw similarity measure summarized above for each song pair. For example, if a rock song and a folk song had a raw similarity of 0.75, and the average similarity between "rock" and "folk" in year x is 0.8, then that genre-weighted similarity between those two songs would be 0.75 * 0.8 = 0.6.[11] If both songs were categorized as "rock," then the weight would equal 1, and the genre-weighted similarity between songs would be 0.75. We then calculated the weighted average of each cell in $\mathbf{V}_{it}$ to create the variable used in our main models*:* a weighted average of each song's distance from all other songs that appeared on the charts in a given year. A simple frequency histogram of this measure provides evidence of the relatively high degree of similarity between songs across our dataset and in popular music more generally ($\mu = 0.81$; $\sigma = 0.06$; Range = 0.26–0.92; see **Figure 1**).[12]

[Figure 1 around here]

Finally, in addition to our yearly genre-weighted typicality measure, we constructed a second variable, *genre-weighted typicality (weekly)*, to investigate week-to-week competition between songs, which we test in our final set of models as a robustness check. Rather than calculating a single typicality score for each song based on its similarity to songs that charted during the 52 weeks prior to its chart debut, we calculated a unique typicality score for each week that a song appears on the charts. To do this, we first measured the cosine similarity between each song's features and those of other songs with which it shared a chart. For each week, we created a matrix $\mathbf{A}_t$ that has dimensions matching the number of songs on each week's charts (100x100), with cell A_{ijt} representing the similarity between song i and song j for that week. Because every song is perfectly similar to itself, we removed **A**'s diagonal from all calculations. As with our yearly typicality measure, we again weighted each cell in **A** by the similarity of each song-pair's genres from the year in which those songs were released. Once these weights were applied, we took the average of each row to give each song-week a *genre-weighted typicality (weekly)* value. This measure is designed to capture how similar a song is to those other songs with which it is directly competing on the charts.

Additional Control Variables. We collected a handful of control variables to account for the multifaceted nature of musical production and ensure the robustness of our effects. First, we included a dummy variable coded to 1 if a song was released on a major record label, and 0 if it was from an independent label. Major labels typically have larger marketing budgets, higher production quality, closer ties with radio stations (see Rossman 2012), and bigger stars on their artist rosters. These factors suggest that songs released by major labels will not only appear more regularly on the charts (two-thirds of the songs in our dataset are major label releases), but that major label releases should have a comparatively easier time reaching the top of the charts. We include the major label dummy in all analyses to account for the benefits that such songs receive when striving to hit the top of the charts.

Second, we included a set of dummy variables in each of our models to account for the number of songs an artist had previously placed on the charts. Musicians receive different levels of institutional support (e.g., marketing or PR), which can affect their opportunities for success, but these differences are difficult to ascertain. These previous song count dummies capture artists' relative visibility or popularity at the moment of a song's release: (1) if a song is an artist's first on the charts, (2) if it is her second or third song on the charts, (3) if it is her fourth through tenth song on the charts, or (4) is she has had more than ten songs in the Hot 100. These dummies also help to capture "superstar" effects (Krueger 2005), which could account for the cumulative advantage popular artists experience as their songs become more likely to climb to the top of the charts.

We also constructed a variable called *multiple memberships* to account for artists who released songs under different names or band formations. For example, Annie Lennox appears on the charts both as a member of the Eurythmics and as a solo artist. As the Eurythmics represent Lennox's first appearance on the charts, every subsequent appearance of hers as a solo artist was coded as a 1 for *multiple memberships*. This was done for every artist who appeared with multiple bands (or with a band and as a solo artist) on the Hot 100 (roughly 6% of our dataset). For these artists, song counts were continued from previous chart incarnations—meaning that Lennox's first charting song as a solo artist was coded as her 15th song overall, because the Eurythmics charted 14 songs before she released her first solo hit. Whether a function of artists' skill in creating chart-friendly songs, labels' commitment to already established artists, or fans' loyalty to certain musicians, maintaining a comprehensive count of previous songs on the charts helps us to account for any potential benefits chart veterans receive.

Fourth, we included a variable called *long song*, set to 1 if a song was unusually long, 0 otherwise. Historical recording formats, along with radio, have encouraged artists to produce songs that are shorter in length, typically between three and four minutes long (Katz 2010). Although the average length of a song on the charts has increased over time, longer songs were likely to get cut short or have trouble finding radio airtime during much of the timeframe covered by our data. We include this dummy to account for the possibility that these difficulties impact chart performance. For our analysis, any song that was two standard deviations longer than the average song for the year in which it was released was denoted a *long song*.

Fifth, we account for "crossover" songs—that is, songs affiliated with multiple genres, and thus (potentially) appealing to multiple audiences.[13] In addition to the Hot 100, *Billboard* has several other, predominantly genre-based charts to capture songs' popularity: mainstream rock, R&B, country, and others. Songs that cross genres and fandom boundaries may be more likely to succeed on the generalist chart (Brackett 1994, Lena 2012), although one could also argue that difficult-to-classify songs may suffer as the result of audience confusion (see Pontikes 2012; Zuckerman 1999). To capture the potential effects of genre-spanning, we created a variable *crossover.* This dummy is coded 1 for any song with more than one song-level genre designation (e.g., blues and country), *unless the two genre designations are pop and rock*, which for much of the chart's history were considered interchangeable and too mainstream to classify across multiple distinct fan bases. *Crossover* is coded as 0 for songs with only one genre classification. Using this method, roughly 24% of the songs in our data are considered crossovers, and on average they perform slightly better on the Hot 100 charts (average peak chart position of 43 versus 45 for crossovers and non-crossovers, respectively; t-test: -3.636, p = .0001).

Sixth, we construct a dummy variable *reissue* for any song that was re-released and appeared on the charts for a second time. As an example, Prince's track "1999" originally charted in 1982, reaching #12 and staying on the charts for 27 weeks. It was reissued for New Years in 1999, when it charted again for a week. Such songs, already familiar to audiences and likely reissued due to their initial popularity, should have an easier time performing well on the charts when they re-enter them. To account for this potential advantage, we coded any song that was re-released in this manner as a *reissue*, and included the dummy in all analyses.

Finally, we included nonparametric time dummies to account for historical variation in our results, partitioning 59 years of data into five-year blocks. This was done for two reasons. First, we wanted to capture the fact that producer and consumer tastes, as well as the sounds and boundaries of certain genres, change over time. Second, we needed a way to account for changes in the underlying calculation and meaning of chart rankings, particularly before and after the move to use SoundScan data (see endnote 6 and Appendix B for further details and analyses). Employing half-decade dummies allows us to estimate and control for the effects of these changes, which had an immediate impact on chart dynamics but took time to be fully understood and absorbed by industry stakeholders (Anand and Peterson 2000). **Table 2** summarizes descriptive statistics and correlations for all the key variables in our analysis.

[Table 2 around here]

Estimation Strategy

To demonstrate the relationship between songs' sonic features and their performance on the Hot 100 charts, we first ran pooled, cross-sectional OLS regressions for each of our two static outcome variables, *peak chart position (inverted)* and *weeks on charts* (Models 1 & 2). These models, run on the 25,102 songs for which we have complete data, are intended to provide correlational face validity of a relationship between our sonic features and chart outcomes.

To conduct a more formal test of the relationship between song typicality and chart performance, and to account for the fact that our *peak chart position* outcome variable is comprised of discrete whole numbers derived from ranks, we run a second set of models using an ordered logit specification (Models 3 & 4). We include various artist-level control variables (e.g., previous song and multiple band membership dummies) instead of artist fixed effects, as ordered logit models with fixed effects can have inconsistent estimators (see Baetschmann, Staub, and Winkelmann 2015). Models estimating *weeks on charts* contain the same control variables, but use a truncated negative binomial specification, as the outcome is a count variable with a minimum value of one (Models 5 & 6).

The models described above reflect cross-sectional analyses that use a song's typicality when it first appeared on the charts to predict its overall success. We know, however, that the Hot 100 charts are dynamic: they are released weekly and change just as frequently, with potentially dozens of songs entering, exiting, and shifting positions. Songs move an average of seven ranks from one week to the next, and they tend to have a relatively short shelf life in the spotlight, with an average chart lifespan of only 11.5 weeks. Following the logic of our earlier prediction, we believe that songs' weekly chart movements will also be influenced by their sonic differentiation from the competition on the Hot 100 charts. Thus, our final set of models estimates the dynamic effect of typicality on inter-song competition.

To conduct this analysis, we model the *weekly* change in songs' chart position as a function of their *genre-weighted typicality (weekly)*. Note that this measure changes as new songs cycle in and out of the charts week-to-week. These models (7 & 8) include linear and quadratic control variables for the number of weeks a song has already been on the charts, as well as song-level fixed effects, which allow us to control for the time-invariant factors of each song, including the artist, the record label, the marketing budget, the song's individual sonic features, and all artist- and song-level controls included in Models 3–6. All independent and control variables are lagged one week—both to (1) match the "natural" one week window used by *Billboard,* and (2) account for the constant short-term churn within the charts, which would render longer lags substantively meaningless. These and all other models presented in the paper employ robust standard errors.

RESULTS

Before presenting our main results, we first wanted to explore the historical relationship between song typicality and chart position. A descriptive analysis of this relationship is presented in **Figures 2a and 2b**, and indicates substantial evolution in the typicality of charting songs over the life-course of our data. To construct these graphs, we took the average typicality of songs during their first week on the charts, and then compared over time (a) those songs that reached the top 40 with those that did not, and (b) those songs that reached number one with those that did not. **Figure 2a** indicates that the songs that peaked in the Top 40 are comparably typical to songs that failed to reach Top 40 status. In fact, in the early years of the charts, top 40 songs are slightly *more* typical than the songs that peaked in positions 41–100. Conversely, **Figure 2b** indicates that, aside from a few punctuated years in which the average number one hit was more typical than the average song on the charts, the most successful songs tended to be *less* typical than other songs, although that gap has narrowed in recent years. Although the average typicality of number one songs is significantly different from that of their peers, they remain close enough to provide *prima facie* support for our optimal differentiation hypothesis.

[Figure 2a and 2b around here]

It is also worth noting the general trends of song typicality across our dataset: the chart's early history was marked by more homogenous, "typical" songs, while more atypical songs tended to appear in the 1970s, '80s, and '90s. This trend toward greater atypicality has reversed itself in recent years, as songs appearing on the charts after 2000 seem to be growing more typical. While these trends tell us something interesting about *absolute* levels of feature-based typicality over time, the models that follow allow us to measure how a song's typicality *relative to its contemporaneous competition* affects its performance on the charts.

Results from our first pair of formal models are depicted in **Figure 3**, which graphically presents standardized estimates of the relationship between songs' sonic features, artists' previous success, and chart performance.[14] These results provide preliminary evidence that some of the sonic features in our dataset are significantly correlated with songs' chart performance, above and beyond the effects of genre, artist, and label affiliation. In Model 1 (represented with white circles), we find that a song's danceability, liveness, and the presence of a 4/4 time signature (as opposed to all other time signatures) are positively associated with peak chart position, while energy (intensity/noise), speechiness, and acousticness produce negative coefficients. Although we do not have space to theorize the interpretation of these individual results, they provide some face validity that product features matter for songs' chart performance.

[Figure 3 around here]

In addition to providing controls for social- and status-related effects on songs' chart position, the dummies for artists' song count reveal evidence of a "sophomore slump." This term refers to the common perception that musicians' often fail to produce a second song or album as popular as their first. Our results provide supporting evidence for this, as an artist's second and third "hit" songs do not perform as well as their first. However, the positive coefficient for songs released by artists with more than 10 previous hits provides support for the "superstar" effects we anticipated, suggesting that artists receive additional advantage after they have achieved substantial popular success.

In Model 2 (represented with black circles), we estimate the effect of these same variables on songs' longevity on the charts (in weeks). These results suggest a similar pattern of relationships, although one difference is worth noting: while we again find evidence of a "sophomore slump," this effect does not reverse as an artist's number of previous hits increases. In other words, if an artist has already charted four or more songs, then subsequent hits will be more likely to experience shorter chart lives. Audiences may more quickly grow tired of music released by artists they already know.

Results from Models 1 and 2 are instructive and provide *prima facie* evidence that sonic features are meaningfully correlated with chart performance. Nevertheless, as they appear independently in these models, the results reveal little about how bundles of features—i.e., songs—are similar to or different from each other *en masse*, or how such differentiation affects chart performance. To address these questions, we move to our next set of models, which use our typicality measure to test how songs' differentiation across feature space affects their performance on the charts. We first discuss the results for our typicality variable of interest across models before examining key control variables.

Table 3 summarizes the coefficients for our key independent and control variables from Models 3–6 (see Appendix **Table A5** for full output). Recall that Model 3 employs ordered logit regression to predict a song's peak position using its typicality relative to other songs that appeared on the charts in the previous 52 weeks. Results suggest a significant negative relationship between song typicality and peak position: controlling for genre affiliation, artist popularity, and a host of other song- and artist-level variables, songs that are more similar to their peers are less likely to reach the top of the charts. In Model 4, we add a squared typicality term to test for our hypothesized inverted U-shaped relationship between typicality and chart performance. Results support our prediction, revealing the benefits of optimal differentiation. The most atypical songs in our dataset would benefit from being more similar to their peers, but as songs become more and more similar, this relationship is reversed—exhibiting too much typicality is associated with poorer chart performance.

[Table 3 about here]

Because second order terms in ordered logit models are difficult to interpret (Karaca-Mandic, Norton, and Dowd 2012), we created **Figure 4** to visualize the marginal effects of songs' typicality on their peak chart position. For purposes of clarity and interpretability we partitioned peak chart position into meaningful "tranches" that are represented by the different lines in the figure. We then use the coefficients from Model 4 to calculate the marginal probability of songs with different typicality levels reaching certain peak positions. Moving from the top of the figure to the bottom (i.e., from the worst position on the charts to the best), we find that the most atypical and most typical songs are likely to fall *outside* of the Top 40 (the white and black circles). These two curves do not reflect the inverted-U shape that we find in Model 4 across the entirety of our dataset, but this makes sense: songs that sound too much (or not enough) like their peers have a higher likelihood of staying outside the top of the charts. The remaining curves—which predict likelihoods of reaching the Top 40, top 20, top 10, top 5, and #1, respectively—all show the expected inverted-U shape relationship, albeit with decreasing likelihoods as each echelon becomes more difficult (and unlikely) for songs to reach. The songs that climb to the top of the charts have a higher marginal probability of doing so when they are in the middle of the typicality distribution—that is, when they are optimally distinct.

As an example of what constitutes an optimally differentiated song, The Beatles' 1969 hit "Come Together" reached the top of the charts on November 29, 1969, and featured a typicality score of 0.66 the week it debuted—over two standard deviations less typical than the average song released that year. Digging into the song's individual features, we find that much of its novelty can be attributed to its low energy (1.2σ below the mean) and low valence (1.9σ below the mean). Although this example does not statistically represent our entire dataset, it does speak to some of the factors that drive our typicality measure and song differentiation in general.

[Figure 4 around here]

Models 5 and 6 employ truncated negative binomial regression to estimate the effect of song typicality on chart longevity. When entered as a linear term, typicality is again negatively associated with length of stay on the charts, but when we include the squared term (Model 6), we once more find an inverted U-shaped relationship. For the most novel songs in our dataset, higher levels of typicality would increase their odds of remaining on the charts, while the most typical songs would remain in the spotlight longer if they were more differentiated. *Ceteris paribus*, a song that is a single standard deviation below mean typicality (0.75 vs. 0.81) is likely to remain on the charts for roughly a half week longer than a song at the mean (11.5 weeks and 11 weeks, respectively).

Across Models 3–6, we find that songs are more likely to attract and maintain the attention of consumers if they are differentiated from other songs on the charts, but not so much that they fail to meet prevailing expectations. We also find consistent results for several of our key control variables. For example, songs released by major labels tend to reach higher chart positions and to last longer on the charts, as we anticipated. Somewhat surprisingly, however, we find that song length is positively related to chart performance (Models 3 and 4). This could be attributed to a few outliers (e.g., Don McLean's "American Pie" is 8:36 long, and spent a month at number 1; The Beatles' "Hey Jude" clocks in at 7:11 and spent 9 weeks at number 1), or it could be evidence of yet another mechanism through which songs achieve some degree of differentiation (although this would not be picked up by our typicality variable). This result seems to indicate that long songs are more salient to listeners than their average-length peers.

As in Models 1 and 2, we again find support for an artist's "sophomore slump" and for "superstar" effects. When looking at the dummies for artists' previous success (the reference categoryhere is an artist's first song on the chart), we find that artists' second and third songs do not do as well as their chart debuts, while songs released by artists with more than 10 previously charting songs reach higher chart positions than do artists' first songs, but they do not stay on the charts as long. These "superstar" effects on peak performance are further supported by the positive coefficient on *multiple memberships*, which suggests that veteran musicians who have already amassed a following as a solo artist or member of a band are likely to see their songs perform better when they hit the charts under a different moniker. Having a pre-established fan base is surely benefitting those artists who, having already proven themselves capable of producing hits, decide to go solo, form a new band, or join a different band altogether. Similarly, we find that "crossovers" benefit from broader audience appeal: songs that span multiple genres are more likely to climb to the top of the charts, although they do not appear to stay on the charts any longer than their single-genre peers.[15]

Finally, we turn to the half-decade time dummies, where the chart's first several years (1958–1961) comprise the omitted reference group. Recall that the five years before (1987–1991) and after (1992–1996) the introduction of SoundScan are of particular interest (note that 1991 is included in the pre-SoundScan era, as the change took place in November of that year). We find that songs were more likely to perform better on the charts prior to SoundScan, whereas in every period thereafter it has become more difficult to reach the top of the charts. Moreover, results from Models 5 and 6 reveal that songs released in the late 1980s and 1990s remain on the charts longer than they did during the earliest years of the Hot 100. This is especially true for songs released directly after the introduction of SoundScan. These results support Anand and Peterson's (2000) claim about how the shift in chart ranking calculation slowed chart churn.

The results presented thus far support our hypothesis that optimally differentiated songs perform better on the charts in general. They do not, however, allow us to speak to the relationship between song typicality and weekly changes in chart position. To explain these week-to-week dynamics, we turn to the results of our fixed effects models, presented in **Table 4**.

[Table 4 around here]

In model 7, the coefficient for *genre-weighted typicality (weekly)* is negative and significant, indicating that songs sounding more similar to their peers are likely to see their performance suffer in subsequent weeks (recall that all covariates and controls are lagged one week in these models). Controlling for the natural decay that songs typically experience on the charts (i.e., the negative coefficient for *weeks on charts*), a single standard deviation increase in typicality results in a song descending more than an *additional* 0.6 positions each week—which is substantial given the relatively low debut position of most songs on the charts (82). The squared *weeks on charts* coefficient is small but positive, reflecting the ever-diminishing distance that songs can drop as they remain on the charts week after week.

Finally, in Model 8 we add a quadratic term for song typicality and find that, all else equal, more typical songs tend to fare worse on subsequent weeks' charts than those that are optimally differentiated. Indeed, only the most novel songs in our dataset benefit from being more similar to the songs around them, suggesting that some degree of typicality is beneficial for success. More practically, this means that songs from heavily underrepresented genres—or songs from mainstream genres that are particularly unique—benefit from the entrance of similar sounding songs, or "sonic neighbors," on the charts. These songs may serve as a kind of bridge for listeners to compare and reconsider songs that are otherwise distinctive. Conversely, songs that would otherwise be deemed too atypical by audiences may perform better when there are other, even more unusual songs already on the charts. For the majority of observations in our dataset, however, increased levels of typicality suggest a subsequent drop in chart position.

DISCUSSION

These results provide evidence that the features of cultural products affect consumption behavior, both independently and in the way they structure how audiences compare and evaluate products (cf. de Vaan et al. 2015; Lena and Pachucki 2013). Controlling for many of the social and industry-specific factors that contribute to a song's success, we find that listeners' assessments of popular music are shaped in part by the content of songs themselves, perhaps suggesting that consumers are more discerning than we sometimes give them credit for (cf. Salganik et al. 2006). Revisiting our initial question, "what makes popular culture popular?", we can add to the list of explanations: (1) the underlying features of products, and (2) the relative position of those products within feature space. Our empirical proxy for this second explanation—typicality, a concept that can easily be adapted to other domains of cultural analysis—significantly predicts how songs perform on the *Billboard* Hot 100 charts. Specifically, we find that most popular songs suffer a penalty for being too similar to their peers, although this effect is attenuated and even reversed for the most novel songs. These effects extend to songs' overall performance, which we measured using peak chart position and longevity, and week-to-week changes in chart position. Our findings support the prediction that songs that manage the similarity-differentiation (or familiarity-novelty) tradeoff are more likely to achieve success.

While we believe that these findings provide important insights into the consumption dynamics of a multi-billion-dollar industry, we also recognize several important limitations. Although the data we use to measure sonic features is relatively comprehensive and sophisticated, it represents a substantial distillation of a song's musical complexity. Reducing such a high dimensional object into eleven fixed features inevitably simplifies its cultural fingerprint and alters its relationships with other like-objects. As MIR tools improve, so too will our ability to map connections between songs. Our data also does not allow us to account for listeners' idiosyncratic interpretations of features or lyric similarity between songs. Moreover, the bounded nature of the Hot 100—it includes only those songs that achieve enough success to appear on the charts in the first place—raises the issue of selection bias and the generalizability of our conclusions. Based on the discussion in Appendix A, we believe that while the most popular cultural products are slightly more typical *on average*, the difference is not so vast as to circumscribe our findings. The patterns we encounter are likely to extend beyond our sample in the music industry and to the other creative industries as well.

The analyses in this paper allow us to explain why, conditional on entering the charts, certain songs outperform others, suggesting that not all popular culture is created equal. In future research, we hope to conduct more dynamic analyses to better understand the nature and implications of specific idiosyncrasies that appear in our dataset. Carving the chart into distinct segments, estimating effects for different time periods, and identifying scope conditions for the arguments presented in this paper will undoubtedly provide additional insight into the dynamic and historically-contingent nature of our findings.

Additionally, although we provide robust evidence for how musical features affect songs' chart performance, our explanation of evaluation outcomes is limited to characteristics of the production environment. Thus, the analyses presented in this paper do not account for the external consumption environment, making it difficult to identify the cognitive mechanisms that drive listeners' selection decisions. Although we suspect that the patterns of optimal differentiation we find are relevant across empirical domains, it remains unclear whether and how these findings could be extended, or whether the concepts herein can prove fruitful for those interested in the ecological dynamics of products that are firmly outside the creative industries. We expect, however, that the curvilinear relationship between typicality and popularity will carry over to other realms of cultural production, such as art, television, and movies. Even the biggest budget productions are likely to be viewed less favorably than their competition if audiences perceive them to be derivative, or too similar to existing productions. We believe that the continued study of the concepts and measures developed in this paper can be generative in a variety of empirical contexts, and serve as a useful tool for social scientists interested in how product features shape consumption behavior. More generally, we hope scholars continue to try and integrate production- and consumption-side narratives to highlight the interdependencies between these processes and their associated outcomes.

Conclusion

Without denying the important role of social dynamics, we remain convinced of the influence product features have on popular success. Both independently and in concert, the content of cultural products needs to be considered more seriously when investigating success in cultural markets. We have demonstrated how a song's feature-derived position amongst its competition—whether considered over the span of a year or a week—contributes to its success. We hope that this paper, including its data and methodological approach, can serve as a model for more content-driven explorations of large-scale empirical puzzles in cultural sociology and beyond.

To that end, we believe that the ideas presented in this paper make several contributions. First, we import methods traditionally associated with computer science and big data analytics to enhance our understanding of large-scale consumption dynamics and performance outcomes. While these tools necessarily simplify the intrinsic high-dimensionality of culture, they also empower us to generate new insights in historically opaque contexts. Although many new cultural measurement tools originate from advances in computer science and other disciplines, social scientists must critically develop and apply them appropriately and thoughtfully (Bail 2014). Other scholars have mapped meaning structures (Mohr 1994, 1998), charted diffusion patterns (Rossman 2012), and introduced the link between cultural content and consumption behavior (Lena 2006; Jones et al. 2012), but there has been no systematic attempt to theorize and measure how product features influence the emergence and diffusion of consumption patterns. In this paper, we introduce and exploit a rich dataset capable of exploring these dynamics, generating new insights into the world of popular music and cultural markets more broadly.

Second, we measure and test the effects of product features and the associations they generate among audiences. Our conception of feature-similarity space can serve both as a tool to map ecosystems of cultural products, and as a means to understand selection dynamics in markets that require subjective evaluation. We argue that the system of associations between products is theoretically and analytically distinct from—though integrally connected to and mediated through—networks of producers and consumers. In so doing, we raise the possibility that cultural content asserts its own autonomous influence over evaluation outcomes through product crowding and differentiation. This conceptualization of culture is dynamic and will ideally push scholars to continue developing new ways to talk about culture and its consequences. One path forward involves importing the tools of network science to study perceived similarities and associations between cultural products. Although existing research on networks focuses largely on interpersonal or interorganizational ties, substantive relationships exist between all sorts of actors, objects, and ideas (Breiger and Puetz 2015). These relationships serve as conduits for information or signals of quality (Podolny 2001), but also as a spatial metaphor for the way in which markets are structured (Emirbayer 1997). Continuing to redefine what constitutes "nodes" and "edges" might help scholars rethink how cultural objects of all types—including products, practices, and ideas—assert influence or agency, thereby addressing a critical issue in social theory more broadly (e.g., Berger and Luckmann 1966). Such a reconception may also change how scholars think about taste formation, which will no longer reside in a theoretical "black box."

Taking these ideas about culture and agency a step further, the dynamics of optimal differentiation also provide a mechanism to support and explain endogenous cultural change (cf. Kaufman 2004, Lieberson 2000). If optimally differentiated products perform better at time *t*, producers seeking success are likely to try and replicate those products in the future. However, given a growing population of producers trying to match the attributes of successful products, and the inevitable lag between production and consumption, the most popular products released at time *t*+1 are likely to come not from producers who earlier chose a replication strategy, but from those who release products that are now optimally differentiated from the competition at *t*+1. As this pattern continues, popular culture will shift and evolve, with products becoming more (and less) typical over time, just as we see in **Figures 2a** and **2b**. The most successful producers, to paraphrase a well-known saying, will be aiming to produce something for where the cultural context is headed, rather than where it currently resides.

Third, our conceptualization of products and feature space contributes to the literature on categories and market structure (e.g., Kovács and Hannan 2015; Pontikes 2012). While research in this area has explored the origins and consequences of categorical classification on firms and products, our results suggest a more grounded approach may be necessary to fully understand how markets are structured. Combinations of features likely play an integral role in the way products, organizations, and even individuals are perceived and evaluated. In our analysis, we include both product features (sonic attributes) and category labels (genres) to ensure that a computer-driven reduction in complexity did not cause inappropriate interpretation. In future work, we intend to dive even deeper into the interrelationship between features and labels. For example, how do product features help create the categorical structure of musicological genres? To draw an analogy, while research has looked at networks of recipe ingredients on the one hand

(Teng, Lin, and Adamic 2012), and the categorization of food and its consequences for market outcomes on the other (Rao, Monin, and Durand 2003; Kovács and Johnson 2013), integrating these perspectives to explore the relationship between ingredients and the way that food is categorized and evaluated appears to be an obvious next step. We hope our findings encourage category scholars to work toward this integration in the study of music, food, and beyond.

Finally, our findings speak to the inherent difficulty—and folly—in practicing "hit song science" (Dhanaraj and Logan 2005; Pachet and Roy 2008). It is certainly true that a small cabal of writers and producers are responsible for many of the most popular songs in recent years (Seabrook 2015), and artists have more tools and data at their disposal than ever before, providing them with incredibly detailed information about the elements of popular songs, which might in turn help them to craft their own hits (Thompson 2014). Nevertheless, while writing recognizable tunes may become easier with the emergence of these tools, our results suggest that artists trying to reverse engineer a hit song may be neglecting two important points. First, songs that sound too similar to the competition are going to have a more difficult time attracting and holding audience attention. Second, and most importantly, the characteristics of contemporaneous songs will have a significant impact on that song's success. Content *and* context matter. Because a song's reception is partially contingent on how differentiated it is from its peers, and artists cannot precisely forecast or control which songs are released concurrently with their own, the crafting of a hit song should be more art than science.

REFERENCES

Ahn, Yong-Yeol, Sebastian E. Ahnert, James P. Bagrow, and Albert-László Barabási. 2011. "Flavor Network and the Principles of Food Pairing." *Scientific Reports* 1.

Alexander, J. C. and P. Smith. 2002. The Strong Program in Cultural Theory: Elements of a Structural Theory. Pp. 135-150 in *Handbook of Sociological Theory*, edited by J. Turner. New York: Kluwer Academic/Plenum Publishers.

Anderson, J. R. 1991. The Adaptive Nature of Human Categorization. Psychological Review 98(3): 409-429.

Alexander, P. J. 1996. Entropy and Popular Culture: Product Diversity in the Popular Music Recording Industry. *American Sociological Review* 61(1):171–74.

Anand, N., and R.A. Peterson. 2000. When Market Information Constitutes Fields: Sensemaking of Markets in the Commercial Music Industry. *Organization Science* 11(3): 270–284.

Baetschmann, Gregori, Kevin E. Staub, and Rainer Winkelmann. 2015. "Consistent Estimation of the Fixed Effects Ordered Logit Model." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 178(3):685–703.

Bail, Christopher A. 2014. The Cultural Environment: Measuring Culture with Big Data. *Theory and Society* 43: 465–482.

Barroso, A., M. S. Giarratana, S. Reis, and O. Sorenson. 2014. "Crowding, Satiation, and Saturation: The Days of Television Series' Lives: Crowding, Satiation, and Saturation." *Strategic Management Journal*.

Becker, Howard. 1982. *Arts Worlds.* Berkeley, CA: University of California Press.

Berger, Jonah and Gael Le Mens. 2009. How Adoption Speed Affects the Abandonment of Cultural Tastes. *Proceedings of the National Academy of Sciences* 106(20): 8146–8150.

Berger, Peter and Thomas Luckmann. 1966. *The Social Construction of Reality: A Treatise in the Sociology of Knowledge.* Garden City, NY: Doubleday & Company, Inc.

Bielby, William T. and Denise D. Bielby. 1994. "'All Hits Are Flukes': Institutionalized Decision Making and the Rhetoric of Network Prime-Time Program Development." *American Journal of Sociology* 99(5):1287–1313.

Blei, D. M., A.Y. Ng, and M. I. Jordan. 2003. "Latent Dirichlet Allocation." *The Journal of Machine Learning Research* 3:993–1022.

Bothner, Matthew S., J. Kang, and T. E. Stuart. 2007. "Competitive Crowding and Risk Taking in a Tournament: Evidence from NASCAR Racing." *Administrative Science Quarterly* 52(2):208–47.

Bourdieu, Pierre. 1993. The Field of Cultural Production: Essays on Art and Literature. New York: Columbia University Press.

Brackett, David. 1994. The Politics and Practice of “Crossover” in American Popular Music, 1963 to 1965. *The Music Quarterly* 78(4): 774–797.

Brashears, Matthew E. and Eric Quintane. 2015. The Microstructures of Network Recall: How Social Networks are Encoded and Represented in Human Memory. *Social Networks* 41: 113–126.

Breiger, Ronald L. and Kyle Puetz. 2015. “Culture and Networks.” Pp. 557–62 in *International Encyclopedia of Social and Behavioral Sciences*, vol. 5, edited by J. D. Wright. Oxford, UK: Elsevier.

Brewer, M.B. 1991. The social self: On being the same and different at the same time. *Personality and Social Psychology Bulletin* 17: 475–482.

Carroll, Glenn R., Olga M. Khessina, David G. McKendrick. 2010. The Social Lives of Products: Analyzing Product Demography for Management Theory and Practice. *The Academy of Management Annals* 4(1): 157–203.

Caves, R.E. 2000. Creative Industries: Contracts Between Art and Commerce. Cambridge, MA: Harvard University Press.

Cerulo, Karen. 1988. Analyzing Cultural Products: A New Method of Measurement. *Social Science Research* 17: 317-352.

Chan, C., J. Berger, and L. Van Boven. 2012. Identifiable but Not Identical: Combining Social Identity and Uniqueness Motives in Choice. *Journal of Consumer Research* 39(3): 561–573.

DellaPosta, Daniel, Yongren Shi, and Michael W Macy. 2015. “Why Do Liberals Drink Lattes?” *American Journal of Sociology* 120(5):1473–1511.

de Vaan, Mathijs, Balazs Vedres, and David Stark. 2015. “Game Changer: The Topology of Creativity.” *American Journal of Sociology* 120(4):1144–94.

Dhanaraj, R. and B. Logan. 2005. Automatic Prediction of Hit Pop Songs. *Proceedings of the International Society for Music Information Retrieval*: 488-491.

DiMaggio, Paul. 1987. Classification in Art. *American Sociological Review* 52(4):440–55.

DiMaggio, Paul, Manish Nag, and David Blei. 2013. Exploiting Affinities between Topic Modeling and the Sociological Perspective on Culture: Application to Newspaper Coverage of U.S. Government Arts Funding. *Poetics* 41: 570–606.

Douglas, M. and B.C. Isherwood. 1996. *The World of Goods: Towards an Anthropology of Consumption.* New York: Routledge.

Dowd, Timothy J. 2004. Concentration and Diversity Revisited: Production Logics and the U.S. Mainstream Recording Market, 1940-1990. *Social Forces* 82(4): 1411–1455.

Eastman, J.T. and T.F. Pettijohn II. 2014. Gone Country: An Investigation of Billboard Country Songs of the Year across Social and Economic Conditions in the United States. *Psychology of Popular Media Culture*.

Emirbayer, Mustafa. 1997. Manifesto for Relational Sociology. *American Journal of Sociology* 103(2): 281–317.

Evans, James A. 2010. "Industry Induces Academic Science to Know Less about More." *American Journal of Sociology* 116(2):389–452.

Fleming, Lee, Santiago Mingo, and David Chen. 2007. Collaborative Brokerage, Generative Creativity, and Creative Success. *Administrative Quarterly* 52: 443–475.

Fligstein, Neil and Luke Dauter. 2007. The Sociology of Markets. *Annual Review of Sociology* 33(1):105–28.

Foster, Jacob G., Andrey Rzhetsky, and James A. Evans. 2015. "Tradition and Innovation in Scientists' Research Strategies." *American Sociological Review*.

Friberg, A., W. Schoonderwaldt, A. Hedblad, M. Fabiani, and A. Elowsson. 2014. Using Listener-based Perceptual Features as Intermediate Representations in Music Information Retrieval. *Journal of the Acoustical Society of America* 136(4): 1951–1963.

Frith, Simon. 1996. Music and Identity. In S. Hall and P. Du Gay (eds.), *Questions of Cultural Identity*. Pp. 108–125. London: Sage.

Frow, J. 1995. *Cultural Studies and Cultural Value*. Oxford: Clarendon Press.

________. 2006. *Genre: The New Critical Idiom*. New York: Taylor and Francis.

Godart, Frédéric C., Andrew V. Shipilov, and Kim Claes. 2014. "Making the Most of the Revolving Door: The Impact of Outward Personnel Mobility Networks on Organizational Creativity." *Organization Science* 25(2):377–400.

Goldberg, Amir. 2011. Mapping Shared Understandings Using Relational Class Analysis: The Case of the Cultural Omnivore Reexamined. *American Journal of Sociology* 116(5): 1397-1436.

Goldberg, Amir., Michael T. Hannan, and Balázs Kovács. 2015. What Does It Mean to Span Cultural Boundaries. *American Sociological Review* (Forthcoming).

Hadida, Allegre. 2015. Performance in the Creative Industries. In The Oxford Handbook of Creative Industries, edited by C. Jones, M. Lorenzen, and J. Sapsed. New York: Oxford University Press.

Hamlen, William A. 1991. Superstardom in Popular Music: Empirical Evidence. *The Review of Economics and Statistics* 73(4):729.

Hannan, Michael T. and John H. Freeman. 1977. The Population Ecology of Organizations. *American Journal of Sociology* 82(5):929–64.

Heckman, James. 1979. Sample selection bias as a specification error. *Econometrica* 47 (1): 153–61.

Hirsch, Paul M. 1972. Processing Fads and Fashions: An Organization-Set Analysis of Cultural Industry Systems. *American Journal of Sociology* 77(4):639–59.

Holt, F. 2007. Genre in Popular Music. Chicago: University of Chicago Press.

Hsu, Greta. 2006. Jacks of all Trades and Masters of None: Audiences' Reactions to Spanning Genres in Feature Film Production. *Administrative Science Quarterly* 51(3): 420–450.

Hsu, Greta. and Michael T. Hannan. 2005. Identities, genres, and organizational forms. *Organization Science* 16(5): 474–490.

Hsu, Greta, Giacomo Negro, and Fabrizio Perretti. 2012. Hybrids in Hollywood: A Study of the Production and Performance of Genre Spanning Films. *Industrial and Corporate Change* 21: 1427-1450.

Huron, David. 2013. A Psychological Approach to Musical Form: The Habituation–Fluency Theory of Repetition. *Current Musicology* 96:7–35.

Jones, Candace, Massimo Maoret, Felipe G. Massa, and Silviya Svejenova. 2012. Rebels with a Cause: Formation, Contestation, and Expansion of the De Novo Category "Modern Architecture," 1870–1975. *Organization Science* 23(6): 1523–1545.

Kahl, Steven J. 2015. Product Conceptual Systems: Toward a Cognitive Processing Model. Pp. 119–46 in *Cognition and Strategy*, vol. 32, edited by G. Gavetti and W. Ocasio. Emerald Group Publishing Limited.

Karaca-Mandic, Pinar, Edward C. Norton, and Bryan Dowd. 2012. "Interaction Terms in Nonlinear Models." *Health services research* 47(1pt1):255–74.

Katz, Mark. 2010. *Capturing Sound: How Technology Has Changed Music*. Berkeley, CA: University of California Press.

Kaufman, J. 2004. Endogenous Explanation in the Sociology of Culture. *Annual Review of Sociology* 30: 335–357.

Keuschnigg, M. 2015. Product Success in Cultural Markets: The Mediating Role of Familiarity, Peers, and Experts. *Poetics* 51:17–36.

Kovács, Balázs and Michael Hannan. 2015. Conceptual Spaces and the Consequences of Category Spanning. *Sociological Science* 2: 252–86.

Kovács, Balázs and Rebeka Johnson. 2013. Contrasting Alternative Explanations for the Consequences of Category Spanning: A Study of Restaurant Reviews and Menus in San Francisco. *Strategic Organization*.

Kovács, Balázs and Amanda Sharkey. 2009. The Paradox of Publicity: How Awards Can Negatively Affect the Evaluation of Quality. *Administrative Science Quarterly* 59(1): 1–33.

Krueger, Alan B. 2005. The Economics of Real Superstars: The Market for Rock Concerts in the Material World. *Journal of Labor Economics* 23(1):1–30.

Lancaster, K. 1975. Socially Optimal Product Differentiation. *The American Economic Review* 65(4): 567–585.

La Rue, J. 2001. Fundamental Considerations in Style Analysis. *The Journal of Musicology* 18(2): 295-312.

Lena, Jennifer C. 2006. Social Context and Musical Content of Rap Music, 1979-1995. *Social Forces* 85(1): 479–495.

_________. 2012. *Banding Together: How Communities Create Genres in Popular Music*. Princeton: Princeton University Press.

_________. 2015. Genre: Relational Approaches to the Sociology of Music. In *Routledge International Handbook of the Sociology of Art and Culture*, edited by Laurie Hanquinet and Mike Savage. Pp. 149–160. New York: Routledge.

Lena, Jennifer C. and Richard A. Peterson 2008. Classification as Culture: Types and Trajectories of Music Genres. *American Sociological Review* 73(5): 697-718.

Lena, Jennifer C. and M.C. Pachucki. 2013. The Sincerest Form of Flattery: Innovation, Repetition, and Status in an Art Movement. *Poetics* 41: 236–264.

Lieberson. Stanley. 2000. *A Matter of Taste: How Names, Fashions, and Culture Change*. New Haven: Yale University Press.

Lizardo, Omar. 2006. How cultural tastes shape personal networks. *American Sociological Review* 71: 778-807.

________. 2014. Omnivorousness as the Bridging of Cultural Holes: A Measurement Strategy. *Theory and Society* 43: 395–419.

Lopes, Paul D. 1992. Innovation and Diversity in the Popular Music Industry, 1969 to 1990. *American Sociological Review* 57(1): 56–71.

Lounsbury, Michael and Mary Ann Glynn. 2001. “Cultural Entrepreneurship: Stories, Legitimacy, and the Acquisition of Resources.” *Strategic management journal* 22(6-7):545–64.

Mark, Noah. 1998. Birds of a Feather Sing Together. *Social Forces* 77(2): 453–485.

_______. 2003. Culture and Competition: Homophily and Distancing Explanations for Cultural Niches. *American Sociological Review* 68(3): 319–345.

Meng, Tao and Mei-Ling Shyu. 2012. Leveraging Concept Association Network for Multimedia Rare Concept Mining and Retrieval. *IEEE International Conference on Multimedia and Expo (ICME) Proceedings*. 860–865.

Mohr, John W. 1994. Soldiers, Mothers, Tramps, and Others: Discourse Roles in the 1907 New York City Charity Directory. *Poetics* 22: 327–357.

________. 1998. Measuring Meaning Structures. *Annual Review of Sociology* 24: 345–370.

Mood, Carina. 2010. “Logistic Regression: Why We Cannot Do What We Think We Can Do, and What We Can Do About It.” *European Sociological Review* 26(1):67–82.

Negus, K. 1992. *Producing Pop: Culture and Conflict in the Popular Music Industry*. London: Edward Arnold.

Nunes, Joseph C. and A. Ordanini. 2014. I Like the Way it Sounds: The Influence of Instrumentation on Pop Song's Place in the Charts. *Musicae Scientiae*: 1–18.

Pachucki, Mark A. and R.L. Breiger. 2010. Cultural Holes: Beyond Relationality in Social Networks and Culture. *Annual Review of Sociology* 36: 205-224.

Pachet, F. and P. Roy 2008. Hit Song Science is Not Yet a Science. *Proceedings of the International Society for Music Information Retrieval*: 355-360.

Peterson, Richard A. 1990. Why 1955?: Explaining the Advent of Rock Music. *Popular Music* 9(1): 97–116.

___________. 1997. *Creating Country Music: Fabricating Authenticity*. Chicago: University of Chicago Press.

Peterson, Richard A. and David C. Berger. 1975. Cycles in Symbol Production: The Case of Popular Music. *American Sociological Review* 40(2): 158–173.

Phillips, Damon J. 2011. Jazz and the Disconnected: City Structural Disconnectedness and the Emergence of a Jazz Canon, 1897–1933. *American Journal of Sociology* 117(2): 420–483.

______________. 2013. *Shaping Jazz: Cities, Labels, and the Global Emergence of an Art Form.* Princeton, NJ: Princeton University Press.

Phillips, Damon J. and Young-Kyu Kim. 2008. Why Pseudonyms? Deception as Identity Preservation Among Jazz Record Companies, 1920-1929. *Organization Science* 20(3):481–99.

Podolny, Joel M. 2001. Networks as the Pipes and Prisms of the Market. *American Journal of Sociology* 107(1):33–60.

Podolny, Joel M., Toby E. Stuart, and Michael T. Hannan.1996. Networks, Knowledge, and Niches: Competition in the Worldwide Semiconductor Industry. *American Journal of Sociology* 102(3): 659–689.

Pontikes, Elizabeth G. 2012. Two Sides of the Same Coin: How Ambiguous Classification Affects Multiple Audiences' Evaluations. *Administrative Science Quarterly* 57(1):81–118.

Pontikes, Elizabeth G. and Michael T. Hannan. 2014. An Ecology of Social Categories." *Sociological Science* 1:311–43.

Rao, Hayagreeva, Philippe Monin, and Rodolphe Durand. 2003. Institutional Change in Toque Vile: Nouvelle Cuisine as an Identity Movement in French Gastronomy. *American Journal of Sociology* 108(4):795–843.

Rosen, Sherwin. 1981. The Economics of Superstars. *The American Economic Review* 71(5): 845-858.

Rossman, Gabriel. 2012. *Climbing the Charts: What Radio Airplay Tells Us about the Diffusion of Innovation.* Princeton, NJ: Princeton University Press.

Rossman, Gabriel. and Oliver Schilke. 2014. Close, But No Cigar The Bimodal Rewards to Prize-Seeking. *American Sociological Review* 79(1):86–108.

Rubio, F.D. 2012. The Material Production of the Spiral Jetty: A Study of Culture in the Making. *Cultural Sociology* 6(2): 143-161.

Salganik, Matthew J., Peter S. Dodds, and Duncan J. Watts. 2006. Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market. *Science* 311: 854-856.

Salganik, Matthew J. and Duncan J. Watts. 2008. Leading the Herd Astray: An Experimental Study of Self-Fulfilling Prophecies in an Artificial Cultural Market. *Social Psychology Quarterly* 71(4):338–55.

_______________. 2009. "Web-Based Experiments for the Study of Collective Social Dynamics in Cultural Markets." *Topics in Cognitive Science* 1(3):439–68.

Schafer, Juliane and Korbinian Strimmer. 2005. A Shrinkage Approach to Large-Scale Covariance Matrix Estimation and Implications for Functional Genomics. *Statistical Applications in Genetics and Molecular Biology* 4(1).

Seabrook, John. 2015. *The Song Machine: Inside the Hit Factory*. 1st Edition. New York City: W. W. Norton & Company.

Serrà, J., Á. Corral, M. Boguñá, M. Haro, and J.L. Arcos. 2012. Measuring the Evolution of Contemporary Western Popular Music. *Scientific Reports* 2.

Shi, Yuan, Olav Sorenson, and David M. Waguespack. The Stability of Centrality-Based Competitive Advantage. *Sociological Science* 4: 107–122.

Smith, Tammy. 2007. Narrative Boundaries and the Dynamics of Ethnic Conflict and Conciliation. *Poetics* 35(1): 22–46.

Sonnett, J. 2004. Musical Boundaries: Intersections of Form and Content. Poetics 32(3-4): 247-264.

Storey, J. 2006. Cultural Theory and Popular Culture: An Introduction. Upper Saddle River, NJ: Pearson Prentice Hall.

Swaminathan, A. and J. Delacroix. 1991. Differentiation within an Organizational Population: Additional Evidence from the Wine Industry. *The Academy of Management Journal* 34(3): 679–692.

Teng, Chun-Yuen, Yu-Ru Lin, and Lada A. Adamic. 2012. Recipe Recommendation Using Ingredient Networks. Pp. 298–307 in *Proceedings of the 3rd Annual ACM Web Science Conference*. ACM.

Thompson, D. 2014. The Shazam Effect. *The Atlantic Magazine*. December.

Tversky, Amos. 1977. Features of Similarity. Psychological Review 84(4): 327-352.

Uzzi, Brian and Jarrett Spiro. 2005. Collaboration and Creativity: The Small World Problem. *American Journal of Sociology* 111: 447-504.

Uzzi, Brian, Satyam Mukherjee, Michael Stringer, and Ben Jones. 2013. Atypical Combinations and Scientific Impact. *Science* 342: 468–472.

Vaisey, Stephen and Omar Lizardo. 2010. Can Cultural Worldviews Influence Network Composition? *Social Forces* 88(4): 1595-1618.

Whitburn, Joel. 1986. *Joel Whitburn's Pop Memories, 1890-1954: The History of American Popular Music: Compiled from America's Popular Music Charts 1890-1954*. Record Research.

_________. 1991. *Joel Whitburn's Top Pop Singles, 1955-1990: Compiled from Billboard's Pop Singles Charts, 1955-1990*. Record Research.

Wijnberg, N.M. and G. Gemser. 2000. Adding Value to Innovation: Impressionism and the Transformation of the Selection System in Visual Arts. Organization Science 11(3): 323-329.

Yogev, Tamar. 2009. The Social Construction of Quality: Status Dynamics in the Market for Contemporary Art. *Socio-Economic Review* 8(3):511–36.

Zajonc, Robert B. 1968. Attitudinal Effects of Mere Exposure. *Journal of Personality and Social Psychology* 9(2p2):1–27.

Zerubavel, Noam, Peter S. Bearman, Jochen Weber, and Kevin N. Ochsner. 2015. Neural Mechanisms Tracking Popularity in Real-World Social Networks. *Proceedings of the National Academy of Sciences (PNAS)* 112(49): 15072–15077.

Zuckerman, Ezra W. 1999. The Categorical Imperative: Securities Analysts and the Illegitimacy Discount. *American Journal of Sociology* 104(5): 1398–1438.

_________. 2004. Structural Incoherence and Stock Market Activity. *American Sociological Review* 69:405–32.

_________. 2016. Optimal Distinctiveness Revisited: An Integrative Framework for Understanding the Balance between Differentiation and Conformity in Individual and Organizational Identities. In *Handbook of Organizational Identity*. Oxford, UK: Oxford University Press.

TABLES AND FIGURES

Table 1. Echo Nest sonic features

Attribute	Scale	Definition
Acousticness	0-1	Represents the likelihood that the song was recorded solely by acoustic means (as opposed to more electronic / electric means)
Danceability	0-1	Describes how suitable a track is for dancing. This measure includes tempo, regularity of beat, and beat strength.
Energy	0-1	A perceptual measure of intensity throughout the track. Think fast, loud, and noisy (i.e., hard rock) more than dance tracks.
Instrumentalness	0-1	The likelihood that a track is predominantly instrumental. Not necessarily the inverse of speechiness.
Key	0-11 (integers only)	The estimated, overall key of the track, from C through B. We enter key as a series of dummy variables
Liveness	0-1	Detects the presence of the live audience during the recording. Heavily studio-produced tracks score low on this measure.
Mode	0 or 1	Whether the song is in a minor (0) or major (1) key
Speechiness	0-1	Detects the presence of spoken word throughout the track. Sung vocals are not considered spoken word.
Tempo	Beats per minute (BPM)	The overall average tempo of a track.
Time Signature	Beats per bar / measure	Estimated, overall time signature of the track. 4/4 is the most common time signature by far, and is entered as a dummy variable in our analyses.
Valence	0-1	The musical positiveness of the track

Note: This list of features includes all but one of the attributes provided by the Echo Nest's suite of algorithms: loudness. This variable was cut from our final analysis at the suggestion of the company's senior engineer, who explained that loudness is primarily determined by the mastering technology used to make a particular recording, a characteristic that is confounded through radio play and other forms of distribution.

Table 2. Correlations and descriptive statistics for select variables in analyses

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]	[16]	[17]	[18]	[19]	[20]	[21]
[1] Peak chart position (inverted)	1																				
[2] Weeks on charts	0.72	1																			
[3] Genre-weighted typicality (yearly)	-0.04	-0.09	1																		
[4] Genre-weighted typicality (weekly)	-0.03	-0.08	0.99	1																	
[5] Major label dummy	0.07	0.10	-0.06	-0.06	1																
[6] Long song	0.03	0.00	-0.06	-0.06	0.02	1															
[7] Crossover track	0.02	0.03	-0.04	-0.04	0.04	0.00	1														
[8] Multiple memberships	0.05	0.01	-0.01	-0.01	0.03	0.02	0.00	1													
[9] 2-3 Previously charting songs	-0.08	-0.02	0.00	0.00	-0.04	-0.03	-0.01	-0.11	1												
[10] 4-10 Previously charting songs	0.01	-0.01	0.01	0.01	0.05	0.00	-0.01	-0.01	-0.32	1											
[11] > 10 Previously charting songs	0.05	-0.06	0.01	0.01	0.06	0.03	0.02	0.26	-0.32	-0.40	1										
[12] Song tempo	-0.01	-0.02	0.07	0.07	-0.01	-0.02	-0.01	0.00	0.00	-0.01	0.01	1									
[13] Song energy	-0.01	0.05	0.16	0.16	0.01	-0.03	0.00	0.03	0.00	0.00	-0.02	0.16	1								
[14] Song speechiness	-0.01	0.05	-0.10	-0.10	0.00	0.02	0.00	-0.02	0.02	-0.02	-0.04	0.01	0.09	1							
[15] Song acousticness	-0.04	-0.19	0.00	-0.01	-0.08	0.00	-0.02	-0.05	-0.01	0.01	0.01	-0.08	-0.56	-0.11	1						
[16] Minor/Major Mode (0 or 1)	-0.01	-0.06	0.66	0.65	-0.01	-0.02	-0.05	0.00	0.00	0.00	0.01	0.03	-0.07	-0.10	0.13	1					
[17] Song danceability	0.03	0.13	0.12	0.13	-0.02	-0.06	0.07	0.02	0.03	-0.03	-0.05	-0.14	0.16	0.19	-0.28	-0.14	1				
[18] Song valence	0.00	-0.05	0.30	0.30	-0.11	-0.10	0.00	-0.01	0.04	-0.02	-0.06	0.09	0.30	0.03	-0.11	-0.04	0.47	1			
[19] Song instrumentalness	0.00	-0.03	-0.30	-0.30	-0.05	0.02	0.01	0.00	0.02	-0.03	-0.05	0.01	-0.08	-0.07	0.12	-0.02	-0.03	0.03	1		
[20] Song liveness	0.05	0.00	-0.11	-0.11	0.01	0.02	-0.04	0.01	-0.01	0.03	0.00	0.01	0.15	0.08	0.02	0.02	-0.22	-0.07	-0.02	1	
[21] Song time singature = 4/4	0.05	0.08	0.09	0.10	0.03	-0.03	0.00	0.03	0.01	0.00	0.00	-0.03	0.27	0.01	-0.27	-0.05	0.26	0.18	-0.07	0.01	1
Mean	56.27	11.57	0.81	0.81	0.67	0.04	0.24	0.08	0.21	0.29	0.28	119.09	0.59	0.07	0.34	0.74	0.58	0.62	0.08	0.24	0.90
Standard Deviation	30.45	7.79	0.06	0.06	0.47	0.19	0.42	0.27	0.40	0.45	0.45	27.70	0.22	0.08	0.31	0.44	0.16	0.24	0.21	0.22	0.29
Minimum	1.00	1.00	0.26	0.26	0	0	0	0	0	0	0	0	0	0.02	0	0	0	0	0	0.01	0
Maximum	100.00	87.00	0.92	0.92	1.00	1.00	1.00	1.00	1.00	1.00	1.00	242.51	1.00	0.96	1.00	1.00	0.99	1.00	1.00	1.00	1.00

Table 3. Select variables from pooled, cross-sectional ordered logit and negative binomial models predicting *Billboard* Hot 100 peak chart position & longevity, 1958-2016

MODEL:	3	4	5	6
	Ordered Logit	Ordered Logit	Negative Binomial	Negative Binomial
OUTCOME VARIABLE:	Peak Position (inverted)	Peak Position (inverted)	Weeks on Charts	Weeks on Charts
Genre-weighted typicality (yearly)	-2.419**	7.672*	-0.538**	1.791
	(0.429)	(2.987)	(0.150)	(1.051)
Genre-weighted typicality (yearly)^2		-6.805**		-1.570*
		(2.004)		(0.698)
Major label dummy	0.145**	0.145**	0.0246**	0.0245**
	(0.0255)	(0.0255)	(0.00883)	(0.00882)
Long song	0.262**	0.265**	0.0291	0.0290
	(0.0609)	(0.0608)	(0.0193)	(0.0193)
2-3 previously charting songs	-0.306**	-0.306**	-0.138**	-0.138**
	(0.0353)	(0.0353)	(0.0119)	(0.0119)
4-10 previously charting songs	-0.0305	-0.0298	-0.118**	-0.118**
	(0.0331)	(0.0331)	(0.0108)	(0.0108)
10+ previously charting songs	0.0874*	0.0878*	-0.168**	-0.168**
	(0.0347)	(0.0347)	(0.0115)	(0.0115)
Crossover track	0.151**	0.149**	-0.00556	-0.00590
	(0.0303)	(0.0303)	(0.0107)	(0.0107)
Multiple memberships	0.146**	0.147**	0.0554**	0.0559**
	(0.0417)	(0.0417)	(0.0133)	(0.0133)
Reissued track	-0.204*	-0.204*	-0.0812*	-0.0814*
	(0.0923)	(0.0921)	(0.0409)	(0.0409)
Half-Decade Dummies				
1987-1991	0.265**	0.232**	0.440**	0.432**
	(0.0697)	(0.0702)	(0.0217)	(0.0218)
1992-1996	-0.282**	-0.328**	0.567**	0.557**
	(0.0701)	(0.0714)	(0.0239)	(0.0241)
Observations	25,077	25,077	25,077	25,077

Robust standard errors in parentheses
Reference categories for dummy variables: Pop (genre), Independent label, 1st charting song (previously charting songs), Key of E-Flat, and all non-4/4 time signatures.
** p<0.01, * p<0.05

Table 4. Results of fixed effects models predicting *Billboard* Hot 100 songs' weekly change in position, 1958-2016

MODEL:	7	8
OUTCOME VARIABLE:	Change in (Inverted) Chart Position	Change in (Inverted) Chart Position
Genre-weighted typicality (weekly)	-10.84**	37.98*
	(2.323)	(15.43)
Genre-weighted typicality (weekly)^2		-32.44**
		(10.33)
Week (on charts)	-1.941**	-1.941**
	(0.0129)	(0.0129)
Week (on charts)^2	0.0334**	0.0334**
	(0.000478)	(0.000477)
Constant	21.95**	3.805
	(1.873)	(5.848)
Observations	263,715	263,715
R-squared	0.432	0.432

Robust standard errors in parentheses
** p<0.01, * p<0.05, two-tailed test

Figure 1. Distribution of genre-weighted song typicality (yearly)

Note: The slight "dip" in this distribution around ~0.80 reflects the binary (0, 1) nature of one of the sonic features included in our typicality measure: mode. Songs written in major and minor keys are equally typical on average, but the sonic distance between a pair of major and minor songs is likely to be greater than a pair pulled at random, producing the bimodal tendency visualized below.

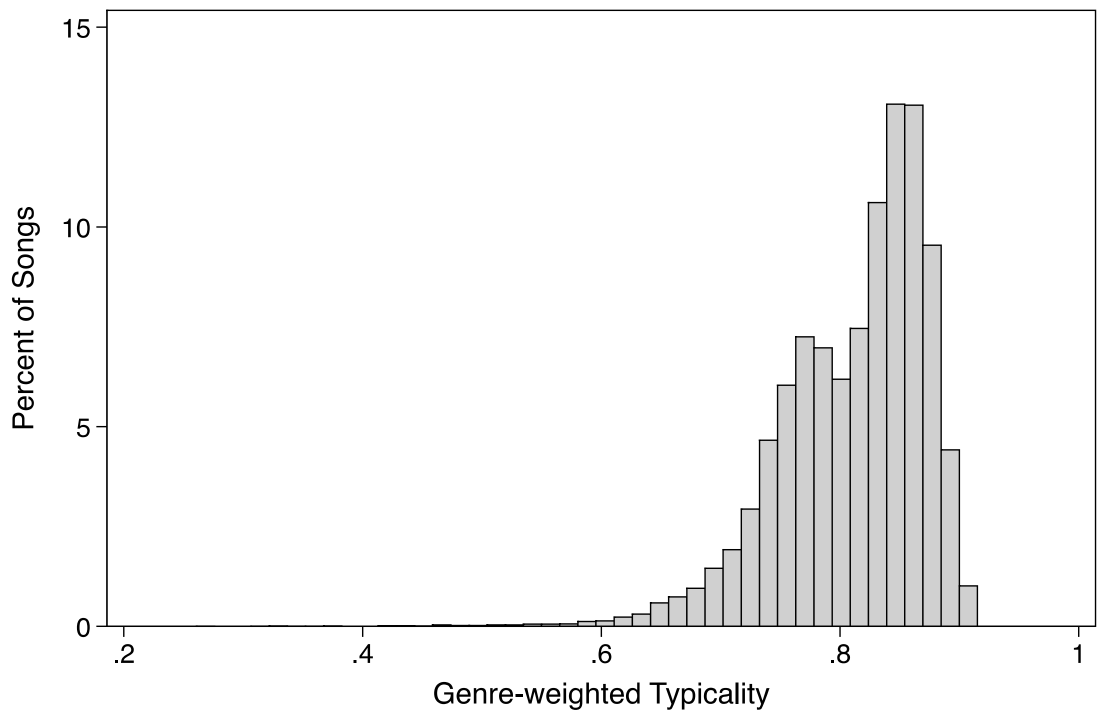


Figure 2a. Comparison of avg. typicality for top 40 songs and all other songs, 1958-2016

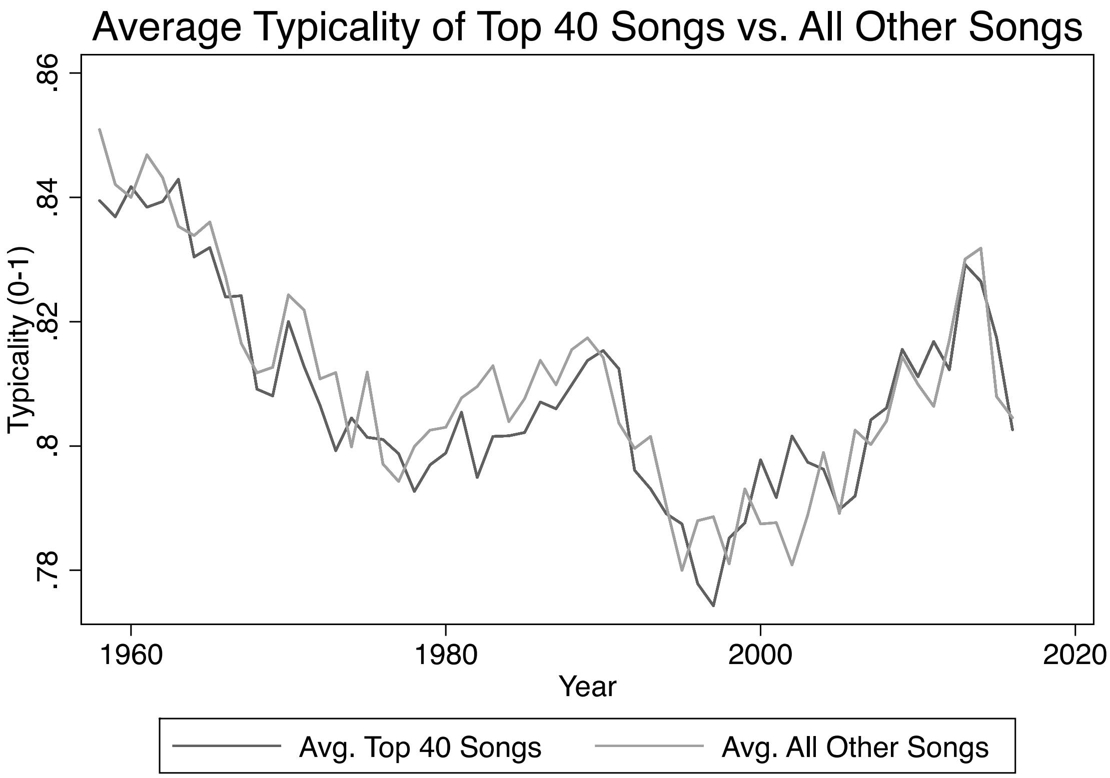


Figure 2b. Comparison of avg. typicality for # 1 songs and all other songs, 1958-2016

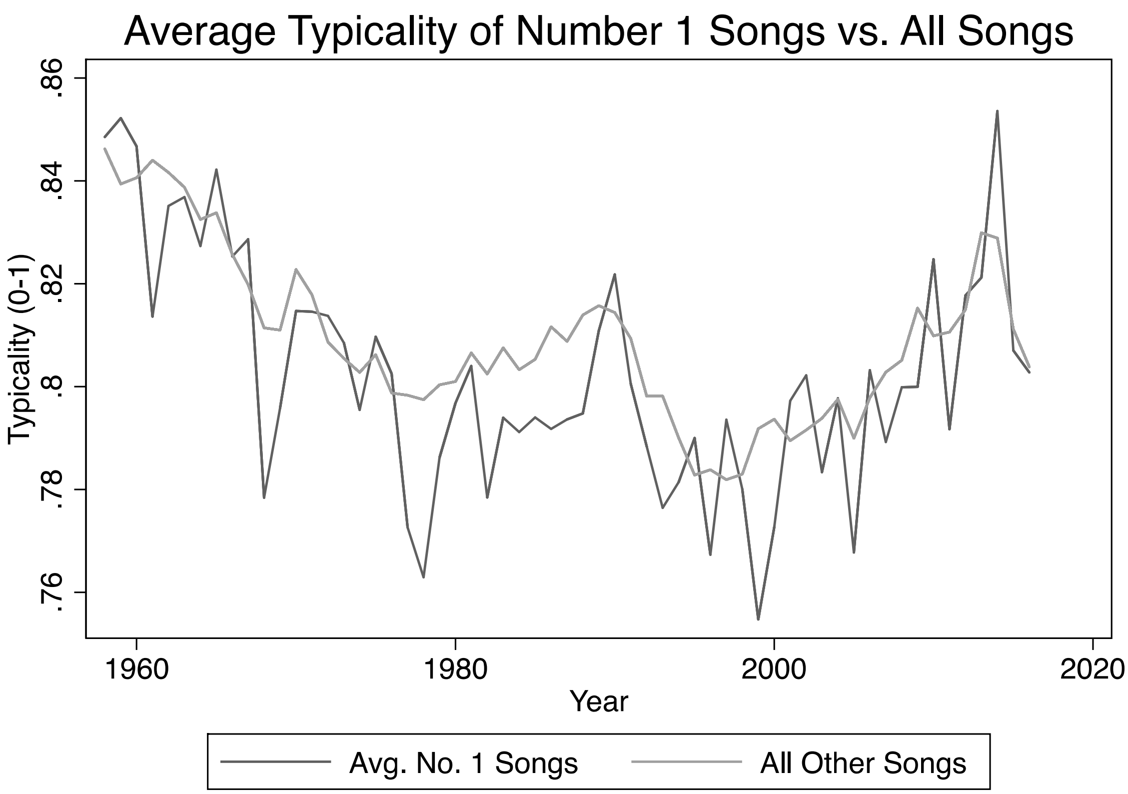


Figure 3. Select standardized coefficients from pooled, cross-sectional OLS models predicting *Billboard* Hot 100 peak chart position & longevity (Models 1 & 2)

Horizontal bars represent 95% CI

See table A4 for full (unstandardized) results.

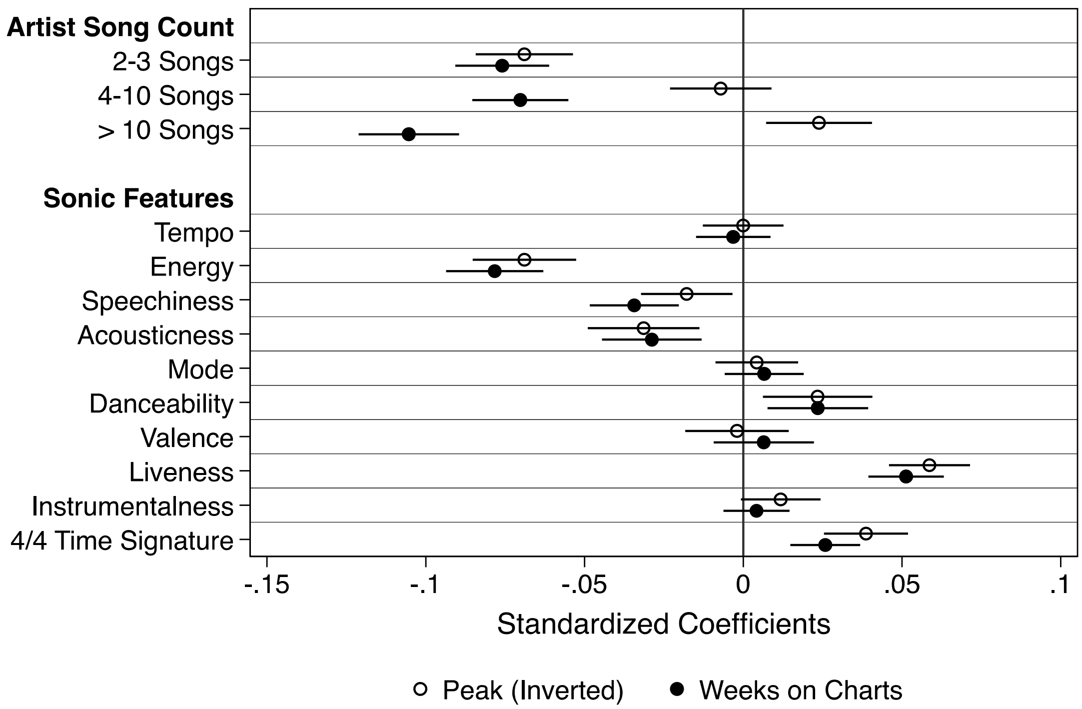
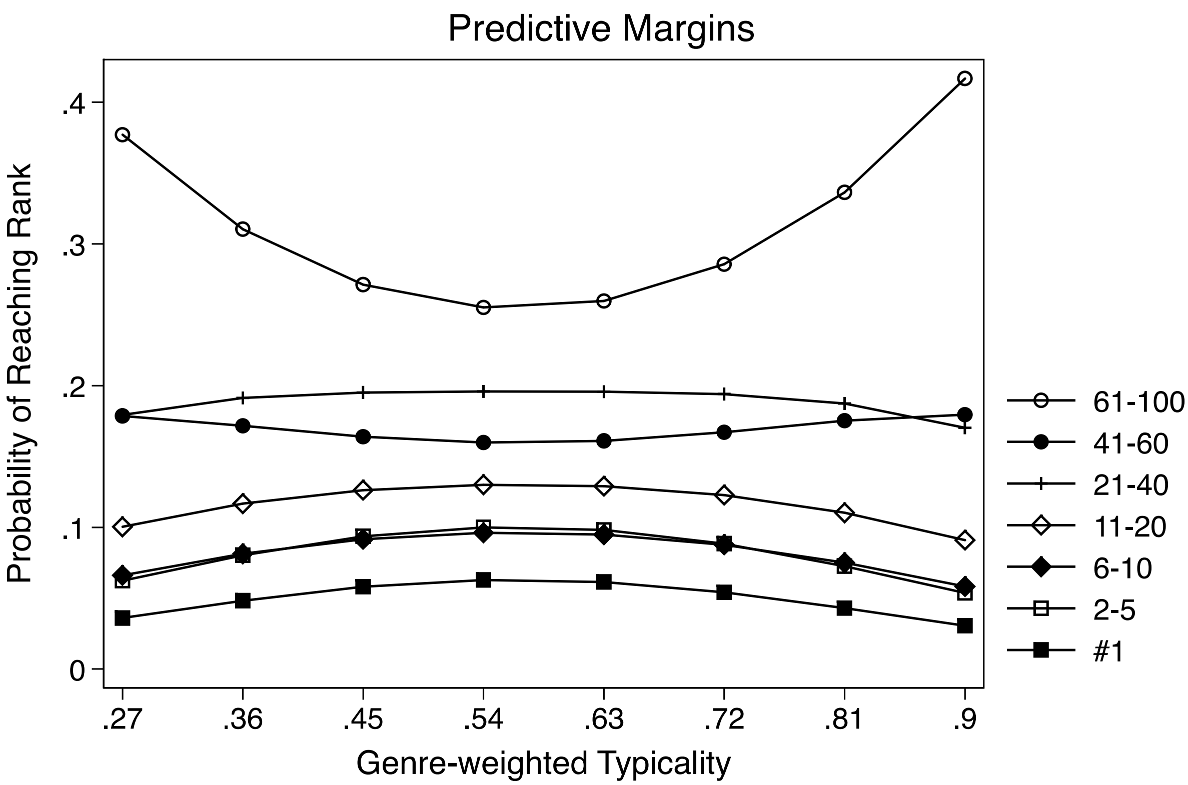


Figure 4. Predicted marginal probability of songs' achieving selected peak position (by typicality) from ordered logit model (Model #4).

Note: Although we inverted chart position in our models to assist readers with a more straightforward interpretation (e.g., positive coefficients reflect better performance), we revert to the originally coded chart positions for our marginal effects graphical analysis. In the figure below, the predicted positions are coded as they would be on the charts (i.e., #100 is the lowest, #1 the highest).



APPENDIX A: Sample Selection Bias and Comparative Analysis

Following previous research that uses chart data to study production and consumption outcomes in music (Lopes 1992; Peterson and Berger 1975), we use data from *Billboard's* Hot 100 charts to answer the question, "what makes popular culture popular?" Nevertheless, we realize that these data constitute a unique and potentially non-representative subset of the world of (Western, popular) music. Without providing additional context or comparison, it is difficult to extrapolate our findings to a broader empirical context outside of similarly unique "best of" lists that sample on the dependent variable. To address this issue, we collected data on two additional sets of songs and conducted a formal comparative analysis that allows us to say more about the generalizability of our results and typicality in music.

To create meaningful samples for comparison, we first noted the genre composition of all Hot 100 songs by year. We then matched the genre-by-year breakdown of the charts with two pseudo-random samples of songs included in the Echo Nest's database (for which we only had access up to 2013): the first sample consists of non-charting songs released by Hot 100 artists, and the second consists of non-charting songs by artists who never appear on the Hot 100. Because of data limitations associated with Discogs (our original source for genre data), we use artist-level genre labels for the comparative analysis. These data come from allmusic.com, another well-established music industry website and data resource. In addition to the more than 25,000 Hot 100 songs analyzed in the body of the paper, we collected sonic feature data on an additional ~40,000 songs: 21,862 non-charting songs by Hot 100 artists, and 18,071 songs by non-Hot 100 artists. As these songs cover roughly the same proportion of genres and years included in our primary analysis, they constitute suitable samples for comparison.

Although the data describing these songs does not allow us to predict an outcome measure equivalent to "chart performance" (these songs never appeared on the *Billboard* charts), we used the Echo Nest's sonic features to calculate typicality scores and other descriptive statistics, which we then compared with the Hot 100 data to assess the representativeness of our original sample. Because we matched songs and artists by genre, we calculated an unweighted typicality score for each song in the two comparison sets. Using the feature data described in the body of the paper, we calculated the typicality score each song by taking the average of its cosine similarity to every other song in the relevant comparison set released that year. The different samples summarized above allow us to create four distinct typicality measures for comparison purposes:

1. ***Hot 100 typicality*** is a baseline typicality measure using just Hot 100 songs. This is the same as the unweighted *all pair typicality (yearly)* variable mentioned in endnote 12.

2. ***Hot 100-other typicality*** includes Hot 100 songs plus non-charting songs by Hot 100 artists (e.g., "unpopular songs by popular artists"). Each song is compared to every other song across *both* subsets in a given year.

3. ***Non-Hot 100 typicality*** combines Hot 100 songs with songs by non-Hot 100 artists (e.g., "songs by unpopular artists"). Again, each song is compared to every other song across both subsets when creating this typicality measure.

4. ***Whole set typicality*** uses all 64,456 songs for which we have data to form a comprehensive comparison set that includes Hot 100 songs, non-charting songs by Hot 100 artists, and songs by non-Hot 100 artists.

These measures allow us to compare how musically similar or homogenous a wider array of songs is—not only to each other, but to songs appearing on the Hot 100 as well. Below, we compare each of these measures and find that while the songs from our primary analysis tend to be slightly more typical than songs that never appeared on the Hot 100 charts, the distributions for each of these measures look remarkably similar.

We include three separate figures to compare these typicality distributions. **Figure A1a** reproduces the original distribution of *Hot 100 typicality*; **Figure A1b** represents the comparative distribution of *Whole set typicality*, grouped by song subset; and **Figure A1c** represents a comparison of the yearly average of *Whole set typicality*, also grouped by song subset.

Figure A1a. Histogram of *Hot 100 typicality*

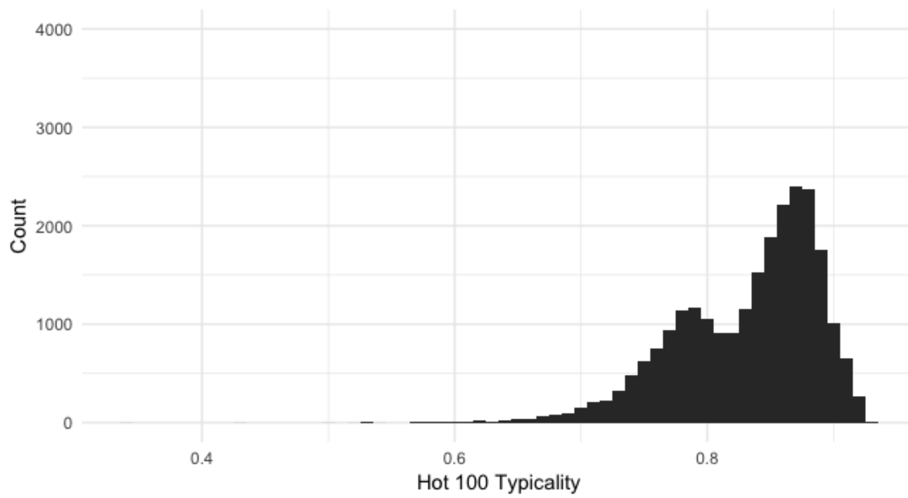


Figure A1b. Histogram of *Whole set typicality*, by song subset

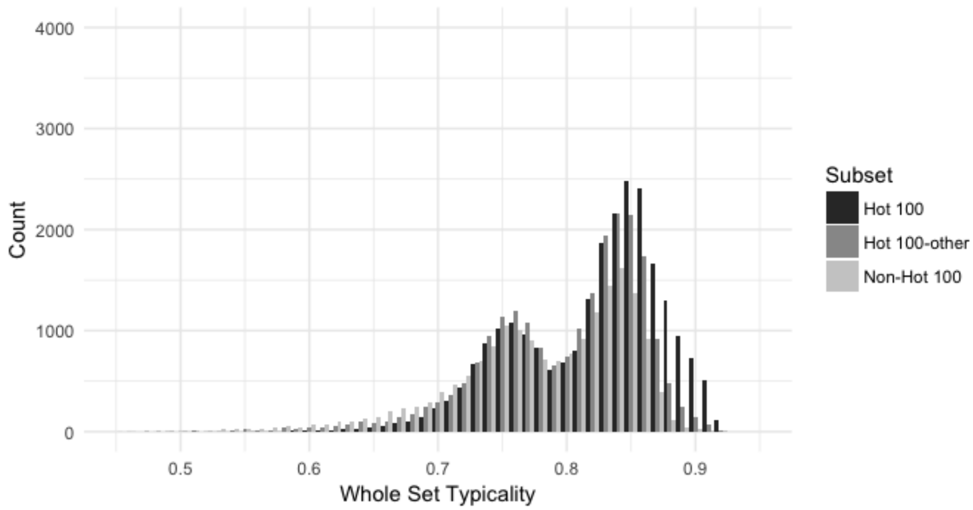
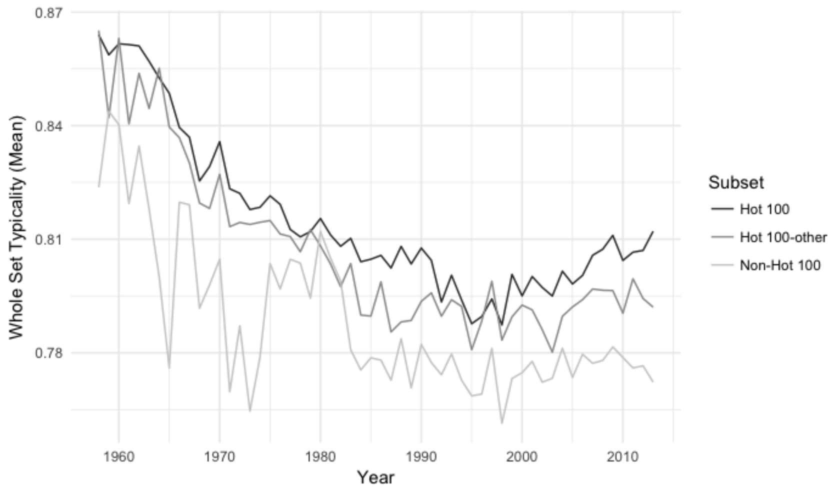


Figure A1c. Average *Whole set typicality* by subset, 1958-2016



In **Figure A1a**, we find that the distribution of *Hot 100 typicality* is nearly identical to that of our primary *Genre-weighted song typicality (yearly)* variable (see **Figure 1**); the two are correlated at R ≈ .95. In **Figure A1b**, we find similarly shaped distributions across each of the three song subsets (*Hot 100*, *Hot 100-other*, and *non-Hot 100*), although songs appearing on the Hot 100 charts are overrepresented at the higher end of the typicality distribution, while songs by artists who never appear on the Hot 100 are overrepresented at the lower end of the distribution. In other words, the most popular songs in our dataset—those that appear on the *Billboard* charts—tend to be slightly more homogenous or typical on average, while songs released by artists who never achieve widespread appeal are likely to be more variable, reflecting greater musical diversity. Finally, in **Figure A1c**, we compare differences in average typicality over time, with Hot 100 songs again appearing at the higher end of the spectrum, but with a similar trend of decreasing typicality over time across each of the subsets.

In **Table A1**, we conduct a more formal comparison of the typicality of Hot 100 songs with each of the other subsets, both individually and collectively. In the first row of the table, we calculate the simple correlation between *Hot 100 typicality* and each of the other measures; all are highly correlated. In the next two rows, we calculate the mean and standard deviation of each typicality measure for Hot 100 songs (row 2) and the relevant comparison group (row 3). While *Whole set typicality* was calculated using all three song subsets, we split each sample out for comparison purposes. Again, we find the different typicality scores for songs on the Hot 100 are fairly stable, and consistently higher than those for both non-charting songs by Hot 100 artists ("Hot 100-other") and songs by artists who never appeared on the charts ("Non-Hot 100"). We also ran two-sample Kolmogorov-Smirnov (K-S) tests to compare the typicality distributions of Hot 100 songs and the comparison groups for the corresponding typicality measure. Lower values indicate distributions that are more similar, while higher values indicate distributions that are farther apart. Although **Figure A1b** suggests that these distributions take on roughly the same shape, a shift in means results in small but significant differences between distributions for each of the typicality measures listed above. Not surprisingly, typicality values for Hot 100 songs are more similar to those for non-charting songs by Hot 100 artists than songs by artists who never appeared on the charts at all.

Table A1. Typicality distribution comparison

	Hot 100-other typicality	**Non-Hot 100 typicality**	**Whole set typicality**	
Correlation with Hot 100 Typicality (for Hot 100 songs)	0.983	0.972	0.969	
Mean (and SD) for songs on Hot 100	0.83 (.057)	0.82 (.057)	0.82 (.057)	
			Hot 100-other	Non-Hot 100
Mean (and SD) for comparison group	0.805 (.063)	0.778 (.069)	0.80 (.062)	0.78 (.069)
Two-sample Kolmogorov-Smirnov (K-S) test statistic	0.145**	0.26**	0.154**	0.257**

**All K-S test statistics significant at $p < .01$, indicating differences in distributions

Finally, to examine whether any systematic bias within the Hot 100 affected our main results, we ran a Heckman two-stage selection model to account for the role typicality plays in getting songs on the chart in the first place (Heckman 1979). In the first-stage probit model, we regressed a binary outcome variable, *In Hot 100*, on *Whole set typicality*, along with artist-level genre and year-level time dummies. The coefficient for *typicality* is positive and significant ($\beta = 2.3$; $p < .001$), indicating that when controlling for genre and time, more typical songs are more likely to appear on the Hot 100 charts (detailed results available upon request). Using this result, we calculated the inverse Mills ratio (IMR) and included it as a control when re-estimating a version of **Model 4** that uses *Whole set typicality* instead of our original genre-weighted measure. The results from the second stage of the Heckman model are presented in **Table A2**. Although our initial sample may suffer from selection bias, it does not alter our findings: the songs that perform best on the Hot 100 are likely to be optimally distinct. This remains true even when typicality is calculated using a much broader set of songs, and when accounting for selection into the charts.

Table A2. Select variables from pooled, cross-sectional ordered logit model predicting *Billboard* Hot 100 peak chart position & longevity, 1958-2013

	Peak Position (Inverted)
Whole set typicality	13.78* (5.786)
Whole set typicality^2	-10.77** (3.625)
Inverse Mills ratio	-0.291** (0.0608)
Major label dummy	0.146** (0.0259)

Long song	0.283**
	(0.0620)
2-3 previously charting songs	-0.301**
	(0.0356)
4-10 previously charting songs	-0.0178
	(0.0334)
10+ previously charting songs	0.0927**
	(0.0353)
Crossover track	0.148**
	(0.0309)
Multiple memberships	0.150**
	(0.0420)
Reissued track	-0.218*
	(0.0998)
Half-Decade Dummies	
1987-1991	0.415**
	(0.0903)
1992-1996	-0.0915
Observations	24,502

Robust standard errors in parentheses
** p<0.01, * p<0.05

These results suggest that popular artists, and popular songs in particular, tend to be more typical and "mainstream" than the field of music writ large. This points to an important boundary condition for our analysis—namely, that the relatively high levels of typicality in the Hot 100 do not fully represent the diversity in music more generally. However, this does not change our main findings. While being overly atypical may cost a song a spot on the charts, exhibiting some degree of musical differentiation can help a song separate itself from the competition and become a "hit."

APPENDIX B: The Introduction of SoundScan and Temporal Variation in our Results

Beginning on November 30, 1991, the *Billboard* Hot 100 implemented a major change in its chart-generating algorithm. Whereas previous charts were created using a combination of record store sales and disc jockey playlists, charts generated from this point forward have relied on data collected systematically by Nielsen's SoundScan. SoundScan automated the collection of sales and airplay data, removing human reporting error from chart calculations. This does not mean that all varieties of human influence were removed from the distribution and sale of popular music (see Rossman 2012), but it does mean that chart data became less susceptible to these biases. *Billboard* has continually changed its methodology—most recently to account for digital downloads and streaming behavior—but previous research suggests that the shift to SoundScan in 1991 had the most substantial effect on the charts (for an in-depth analysis of this shift and its consequences, see Anand and Peterson [2000]).

To test whether and how this shift affects our results, we re-estimated our main models for the years directly preceding and following the introduction of SoundScan. For songs appearing on the Hot 100 between 1986–1990 and 1992–1996,[16] our measure of *genre-weighted typicality (yearly)* fails to significantly predict peak position or chart longevity, although results for the five-year period following the introduction of SoundScan are directionally in line with our main findings. This is also true when using data from the full decades before and after SoundScan (i.e., 1981–1990 and 1992–2001; results available upon request). To assess how meaningful these differences are, we conducted a simple z-test to compare the coefficients on our linear and squared terms of our typicality variable in models 4 and 6, but estimated using only data from 1986–1990 (or 1981–1990) and 1992–1996 (1992–2001). Though comparing coefficients across logistic regression results is not without its concerns (see Mood 2010), in each case we found that the results from these periods are not significantly different from one another. This suggests that although we observe variation in our results around the introduction of SoundScan, such variation falls within the range of what we would expect given heterogeneity in our data.

While the shift to SoundScan is certainly an important event in the history of the *Billboard* charts, these results raise the question of temporal variation in our data more generally. Other possible sources of variation include the fragmentation of rock music into distinct sub-genres in the 1970s; the rise of compact-disc technology in the 1980s; the consolidation of major labels in the 1970s and 1980s; the emergence of indie labels in the 1990s; and the rise of iTunes (and later online streaming) in the 2000s, just to name a few. Indeed, each of these events could act as an exogenous shock that causes significant variation in the typicality of songs appearing on the Hot 100 charts, and of our results more generally. It is not our goal in this paper to identify and explain each of these possible shifts, but we did want to see if the difference we find pre- versus post-SoundScan represents a unique discontinuity, or is instead one of many examples of historical variation in our data.

To assess how remarkable the SoundScan disjuncture was in broader historical context, we computed and compared results across models 4 and 6 from the main text for each five-year period in our dataset (e.g., 1958–1962, 1959–1963, … , 2012–2016; full results available upon request). More specifically, after estimating results for each of these periods, we (1) calculated the difference between coefficients for each sequential five-year period (e.g., *β(1963–1967) – β(1958–1962)*); (2) took the absolute value of each of these differences and listed them in order of magnitude; and then (3) checked to see where the pre- versus post-SoundScan difference fell in this ordered list. Out of all the differences in our results across these rolling five-year periods, the pre- versus post-SoundScan difference falls in the 46th percentile. Put another way, more than half of the five-year periods compared in this back-of-the-envelope analysis generated larger differences than the one we find between 1987–1991 and 1992–1996. This again suggests that the variation produced by Soundscan falls within the range of expected temporal variation.

To further test this assumption and see whether SoundScan's effect on the charts was discontinuous or part of a broader historical trend, we plotted several summary statistics in our data by year, including: (1) the number of distinct songs appearing on the Hot 100 charts (**Figure A2a**); (2) the number of distinct genres appearing on the charts (**Figure A2b**); (3) the count of #1 hits (**Figure A2c**); (4) the average peak position of Hot 100 songs (**Figure A2d**); and (5) the average number of weeks songs stayed on the charts (**Figure A2e**). Although these figures provide some evidence of noticeable discontinuities associated with the introduction of SoundScan—such as a drop in the count of #1 hits per year (from 27 to 12) and the average peak position of songs on the charts (from 35 to 45)—they largely suggest broader historical trends. For example, the decrease in number of unique songs per year between 1992 and 2000 represents the continuation of a trend started in the 1960s. Moreover, while this and other shifts (e.g., the rise in average number of weeks on the charts) seem to affect the period directly following SoundScan (i.e., the 1990s), they reverse course and regress toward the mean in the 2000s.

Figure A2a. Count of Songs Appearing on the Hot 100, per year

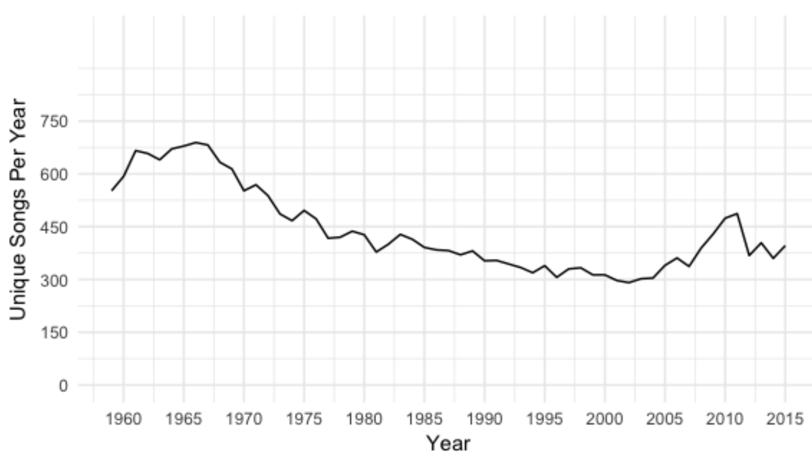


Figure A2b. Count of Genres Appearing on the Hot 100, per year

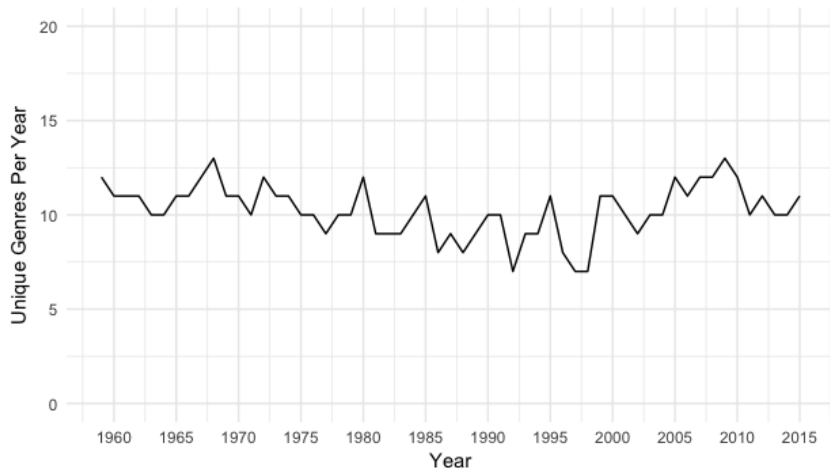


Figure A2c. Count of #1 Hits Appearing on the Hot 100, per year

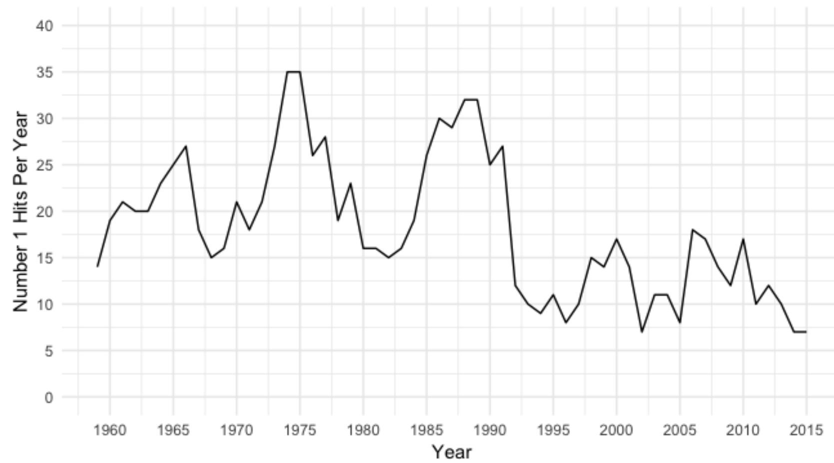


Figure A2d. Average Peak Position of Songs Appearing on the Hot 100, per year

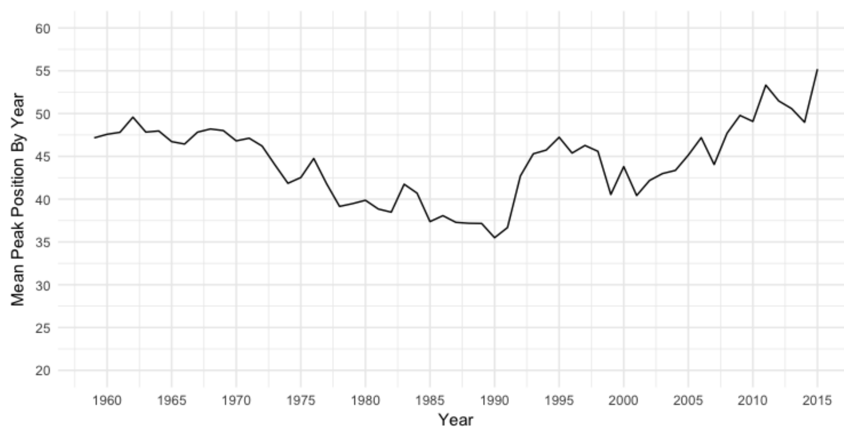
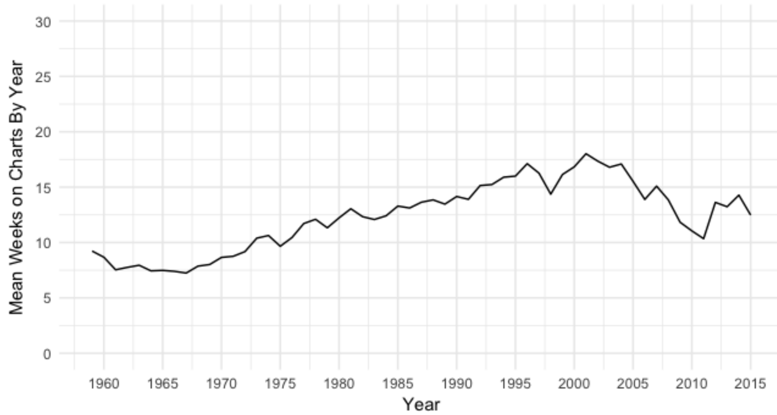


Figure A2e. Average # of Weeks on Chart for Songs Appearing on the Hot 100, per year



Taken together, these analyses suggest that (1) any differences associated with SoundScan fall within the expected range of historical variation in our data, and (2) many of the associated shifts in the makeup of the charts around this time are either temporary or part of a broader historical trend. As shown in **Table A3**, this second point is supported when we estimate our main models using data from 2001–2010. While the period directly following the introduction of SoundScan (1992–2001) fails to reproduce our main results, the subsequent decade exhibits the expected inverted-U-shaped relationship between song typicality and chart performance—both in terms of peak position reached and number of weeks on the charts. It is clear that SoundScan represents an important source of variation in our results, but it is not the only source.

Moreover, while this variation is certainly interesting, we do not believe it calls our main findings into question; rather, it highlights the historical contingency of our results, which is not uncommon in social science studies spanning multiple decades like ours (see Shi, Sorenson, and Waguespack 2017). In future research, we hope to investigate possible explanations for this variation and develop historical boundary conditions for our findings, but it is beyond the scope of this paper to do so. In our primary analyses, we account and control for this temporal variation non-parametrically with our five-year dummy variables.[17]

Table A3. Select variables from pooled, cross-sectional ordered logit and negative binomial models predicting *Billboard* Hot 100 peak chart position & longevity, 2001–2010

MODEL SPECIFICATION:	Ordered Logit	Negative Binomial
OUTCOME VARIABLE:	Peak Position (inverted)	Weeks on Charts
Genre-weighted typicality (yearly)	21.26**	7.815*
	(5.617)	(3.138)
Genre-weighted typicality (yearly)^2	-15.67**	-5.985**
	(4.149)	(2.188)
Major label dummy	0.170*	0.0475
	(0.0681)	(0.0283)
Long song	0.168	-0.00985
	(0.163)	(0.0685)
2-3 previously charting songs	-0.0577	-0.0628
	(0.0986)	(0.0376)
4-10 previously charting songs	0.224*	-0.000876
	(0.0925)	(0.0334)
10+ previously charting songs	-0.0110	-0.257**
	(0.0966)	(0.0386)
Crossover track	0.142	-0.150**
	(0.0779)	(0.0346)
Multiple memberships	0.0527	0.0947
	(0.144)	(0.0525)
Reissued track	-0.126	0.191
Observations	3,474	3,474

Robust standard errors in parentheses.
Coefficients for genre, key, and sonic features not shown but included in models. Year dummies not included for the shorter time periods covered.
** p<0.01, * p<0.05

APPENDIX C: Full Model Results

Table A4. Results from pooled, cross-sectional OLS models predicting *Billboard* Hot 100 peak chart position & longevity (Models 1 & 2 / Figure 3 in main text)

MODEL:	1	2
OUTCOME VARIABLE:	Peak Position (Inverted)	Weeks on Charts
Major label dummy	2.544**	0.298**
	(0.431)	(0.101)
Long song	4.220**	0.165
	(1.007)	(0.228)
2-3 previously charting songs	-5.201**	-1.465**
	(0.589)	(0.145)
4-10 previously charting songs	-0.477	-1.210**
	(0.550)	(0.133)
10+ previously charting songs	1.616**	-1.825**
	(0.576)	(0.140)
Crossover track	2.543**	-0.0812
	(0.510)	(0.124)
Multiple memberships	2.279**	0.550**
	(0.706)	(0.158)
Reissued track	-3.297*	-0.882*
	(1.642)	(0.415)
Genre Dummies		
Blues	-12.17**	-0.734
	(1.954)	(0.383)
Brass & Military	-8.505	-3.422
	(12.84)	(1.760)
Children's	5.264	-3.011
	(7.229)	(1.600)
Classical	-5.536	0.436
	(8.032)	(1.923)
Electronic	1.259	1.238**
	(0.862)	(0.235)
Folk, World, & Country	-8.663**	0.102
	(0.822)	(0.208)
Funk / Soul	-4.299**	0.316*
	(0.721)	(0.155)
Hip Hop	1.472	0.508
	(0.962)	(0.272)
Jazz	-11.98**	-0.678**
	(1.333)	(0.246)
Latin	-18.15**	-3.137**
	(3.430)	(0.832)
Non-Music	0.149	-0.0113
	(4.154)	(0.793)
Reggae	-3.166	0.873
	(3.974)	(1.010)
Rock	1.004	0.672**
	(0.637)	(0.151)
Stage & Screen	-4.968	-0.685
	(4.937)	(1.324)
Sonic Features		
Tempo	-1.33e-05	-0.000887
	(0.00713)	(0.00168)
Energy	-9.710**	-2.822**
	(1.171)	(0.282)
Speechiness	-6.409*	-3.160**
	(2.645)	(0.659)
Acousticness	-3.098**	-0.728**
	(0.886)	(0.202)
Mode (1 = Major key)	0.297	0.118
	(0.464)	(0.113)
Danceability	4.584**	1.176**
	(1.719)	(0.405)
Valence	-0.253	0.213

	(1.073)	(0.266)
Instrumentalness	1.686	0.154
	(0.914)	(0.194)
Liveness	8.100**	1.812**
	(0.898)	(0.214)
Key = C	-0.685	-0.331
	(0.953)	(0.243)
Key = C-sharp	0.705	0.152
	(1.020)	(0.270)
Key = D	-1.232	-0.519*
	(0.985)	(0.249)
Key = E-flat	1.440	0.170
	(1.246)	(0.312)
Key = E	-1.085	-0.495
	(1.039)	(0.264)
Key = F	-0.215	-0.255
	(1.014)	(0.255)
Key = F-sharp	-1.305	-0.396
	(0.961)	(0.246)
Key = G	0.408	-0.213
	(1.096)	(0.278)
Key = G-sharp	0.252	-0.216
	(0.978)	(0.248)
Key = A	-0.132	-0.229
	(1.053)	(0.267)
Key = B-flat	0.194	-0.119
	(1.065)	(0.282)
4/4 time signature dummy	4.010**	0.686**
	(0.700)	(0.149)
Half-Decade Dummies		
1962-1966	-0.624	-0.644**
	(0.885)	(0.140)
1967-1971	-1.549	-0.329*
	(0.924)	(0.154)
1972-1976	1.810	1.592**
	(0.977)	(0.177)
1977-1981	3.809**	3.304**
	(0.996)	(0.206)
1982-1986	2.865**	3.753**
	(1.044)	(0.215)
1987-1991	5.547**	4.856**
	(1.113)	(0.231)
1992-1996	-2.379*	6.881**
	(1.143)	(0.280)
1997-2001	0.638	7.401**
	(1.141)	(0.297)
2002-2006	0.500	7.583**
	(1.155)	(0.302)
2007-2011	-4.455**	4.294**
	(1.101)	(0.305)
2012-2016	-5.970**	5.527**
	(1.224)	(0.380)
Constant	54.82**	9.537**
	(2.015)	(0.460)
Observations	25,102	25,102
R-squared	0.047	0.171

Robust standard errors in parentheses
Reference categories are: Pop (genre), Independent label, 1st charting song (previously charting songs), Key of E-Flat (key), and all time signatures other than 4/4.
** p<0.01, * p<0.05

Table A5. Results from pooled, cross-sectional ordered logit models predicting *Billboard* Hot 100 peak chart position & longevity, 1958-2013 (Table 3 in main text)

MODEL:	3	4	5	6
	Ordered Logit	Ordered Logit	Negative Binomial	Negative Binomial
OUTCOME VARIABLE:	Peak Position (inverted)	Peak Position (inverted)	Weeks on Charts	Weeks on Charts
Genre-weighted typicality (yearly)	-2.419**	7.672*	-0.538**	1.791
	(0.429)	(2.987)	(0.150)	(1.051)
Genre-weighted typicality (yearly)^2		-6.805**		-1.570*
		(2.004)		(0.698)
Major label dummy	0.145**	0.145**	0.0246**	0.0245**
	(0.0255)	(0.0255)	(0.00883)	(0.00882)
Long song	0.262**	0.265**	0.0291	0.0290
	(0.0609)	(0.0608)	(0.0193)	(0.0193)
2-3 previously charting songs	-0.306**	-0.306**	-0.138**	-0.138**
	(0.0353)	(0.0353)	(0.0119)	(0.0119)
4-10 previously charting songs	-0.0305	-0.0298	-0.118**	-0.118**
	(0.0331)	(0.0331)	(0.0108)	(0.0108)
10+ previously charting songs	0.0874*	0.0878*	-0.168**	-0.168**
	(0.0347)	(0.0347)	(0.0115)	(0.0115)
Crossover track	0.151**	0.149**	-0.00556	-0.00590
	(0.0303)	(0.0303)	(0.0107)	(0.0107)
Multiple memberships	0.146**	0.147**	0.0554**	0.0559**
	(0.0417)	(0.0417)	(0.0133)	(0.0133)
Reissued track	-0.204*	-0.204*	-0.0812*	-0.0814*
	(0.0923)	(0.0921)	(0.0409)	(0.0409)
Genre Dummies				
Blues	-0.727**	-0.732**	-0.123**	-0.123**
	(0.111)	(0.111)	(0.0446)	(0.0447)
Brass & Military	-1.173	-0.911	-0.481*	-0.413*
	(1.003)	(0.961)	(0.223)	(0.210)
Children's	-0.157	-0.0278	-0.347	-0.334
	(0.461)	(0.431)	(0.210)	(0.204)
Classical	-0.723	-0.642	-0.0555	-0.0241
	(0.484)	(0.507)	(0.164)	(0.167)
Electronic	0.122*	0.127*	0.100**	0.101**
	(0.0520)	(0.0521)	(0.0190)	(0.0190)
Folk, World, & Country	-0.477**	-0.482**	0.0134	0.0123
	(0.0463)	(0.0464)	(0.0189)	(0.0190)
Funk / Soul	-0.250**	-0.250**	0.0230	0.0229
	(0.0421)	(0.0422)	(0.0148)	(0.0148)
Hip Hop	0.127*	0.130*	0.0458*	0.0467*
	(0.0576)	(0.0576)	(0.0212)	(0.0212)
Jazz	-0.749**	-0.750**	-0.0947**	-0.0947**
	(0.0816)	(0.0815)	(0.0291)	(0.0291)
Latin	-1.240**	-1.224**	-0.353**	-0.343**

	(0.213)	(0.217)	(0.0919)	(0.0926)
Non-Music	-0.298	-0.201	-0.117	-0.0991
	(0.215)	(0.213)	(0.0851)	(0.0836)
Reggae	-0.389	-0.323	0.0442	0.0548
	(0.263)	(0.256)	(0.0764)	(0.0760)
Rock	0.0569	0.0581	0.0638**	0.0640**
	(0.0375)	(0.0375)	(0.0137)	(0.0137)
Stage & Screen	-0.588**	-0.502	-0.113	-0.0806
	(0.278)	(0.280)	(0.139)	(0.141)
Sonic Features				
Tempo	1.94e-05	5.75e-05	-3.24e-05	-2.80e-05
	(0.000419)	(0.000419)	(0.000146)	(0.000146)
Energy	-0.473**	-0.442**	-0.216**	-0.209**
	(0.0717)	(0.0722)	(0.0242)	(0.0244)
Speechiness	-0.524**	-0.539**	-0.252**	-0.258**
	(0.157)	(0.156)	(0.0531)	(0.0531)
Acousticness	-0.187**	-0.170**	-0.0763**	-0.0722**
	(0.0527)	(0.0528)	(0.0178)	(0.0180)
Mode (1 = Major key)	0.246**	0.304**	0.0597**	0.0727**
	(0.0489)	(0.0511)	(0.0173)	(0.0179)
Danceability	0.397**	0.425**	0.138**	0.143**
	(0.104)	(0.104)	(0.0358)	(0.0358)
Valence	0.0844	0.118	0.0324	0.0400
	(0.0666)	(0.0670)	(0.0233)	(0.0236)
Instrumentalness	-0.104	-0.143*	-0.0148	-0.0231
	(0.0644)	(0.0650)	(0.0214)	(0.0216)
Liveness	0.399**	0.378**	0.139**	0.134**
	(0.0549)	(0.0552)	(0.0182)	(0.0184)
Key = C	-0.109*	-0.129*	-0.0383*	-0.0428*
	(0.0544)	(0.0547)	(0.0189)	(0.0190)
Key = C-sharp	0.00510	-0.00605	0.0143	0.0117
	(0.0584)	(0.0585)	(0.0203)	(0.0203)
Key = D	-0.0928	-0.0934	-0.0371	-0.0374
	(0.0552)	(0.0553)	(0.0193)	(0.0193)
Key = E-flat	0.0897	0.100	0.0270	0.0291
	(0.0719)	(0.0720)	(0.0254)	(0.0254)
Key = E	-0.0459	-0.0332	-0.0205	-0.0179
	(0.0586)	(0.0588)	(0.0208)	(0.0208)
Key = F	0.0226	0.0371	-0.00155	0.00175
	(0.0582)	(0.0584)	(0.0200)	(0.0200)
Key = F-sharp	0.0396	0.0557	0.0171	0.0206
	(0.0630)	(0.0632)	(0.0223)	(0.0223)
Key = G	-0.0510	-0.0322	-0.0198	-0.0157
	(0.0548)	(0.0551)	(0.0191)	(0.0191)
Key = G-sharp	0.0459	0.0609	-0.00407	-0.000956
	(0.0630)	(0.0632)	(0.0218)	(0.0218)
Key = A	0.0227	0.0327	-0.00348	-0.00139
	(0.0548)	(0.0549)	(0.0189)	(0.0189)
Key = B-flat	0.000810	0.00501	-0.00452	-0.00398
	(0.0603)	(0.0603)	(0.0207)	(0.0207)
4/4 time signature dummy	0.262**	0.266**	0.0742**	0.0747**

	(0.0414)	(0.0413)	(0.0147)	(0.0147)
Half-Decade Dummies				
1962-1966	-0.0647	-0.0707	-0.0964**	-0.0980**
	(0.0530)	(0.0529)	(0.0177)	(0.0177)
1967-1971	-0.156**	-0.179**	-0.0586**	-0.0637**
	(0.0556)	(0.0559)	(0.0188)	(0.0189)
1972-1976	0.0234	-0.0107	0.157**	0.150**
	(0.0606)	(0.0614)	(0.0201)	(0.0202)
1977-1981	0.0811	0.0418	0.315**	0.307**
	(0.0606)	(0.0616)	(0.0213)	(0.0215)
1982-1986	0.0425	0.00613	0.353**	0.345**
	(0.0638)	(0.0647)	(0.0215)	(0.0216)
1987-1991	0.265**	0.232**	0.440**	0.432**
	(0.0697)	(0.0702)	(0.0217)	(0.0218)
1992-1996	-0.282**	-0.328**	0.567**	0.557**
	(0.0701)	(0.0714)	(0.0239)	(0.0241)
1997-2001	-0.108	-0.156*	0.603**	0.593**
	(0.0709)	(0.0726)	(0.0245)	(0.0247)
2002-2006	-0.0931	-0.136	0.623**	0.614**
	(0.0704)	(0.0716)	(0.0249)	(0.0251)
2007-2011	-0.350**	-0.379**	0.392**	0.385**
	(0.0665)	(0.0670)	(0.0277)	(0.0277)
2012-2016	-0.414**	-0.433**	0.495**	0.492**
	(0.0731)	(0.0734)	(0.0305)	(0.0305)
Observations	25,077	25,077	25,077	25,077

Robust standard errors in parentheses
Reference categories for dummy variables: Pop (genre), Independent label, 1st charting song (previously charting songs), Key of E-Flat, and all non-4/4 time signatures.
** p<0.01, * p<0.05

ENDNOTES

[1] To avoid repetition, we use the terms "features," "attributes," and "characteristics" interchangeably to refer to the fixed, material elements that constitute cultural products. For example, in the context of a painting, relevant features might include the different colors used, along with whether the painting is a portrait or a landscape.

[2] We recognize that category labels might themselves be considered just another product attribute or feature, but we treat them here as distinct entities. This distinction is analytical as well as phenomenological, as category labels convey a qualitatively different kind of information than the underlying features of products.

[3] We use the term "success" in this paper to connote mass or popular appeal, rather than critical acclaim or other legitimate measures of performance.

[4] While we do not explicitly invoke network terminology to describe our theory—in part because we do not use network measures to test it—the notion of a product "association network" can serve as a salient image to help visualize this space. Although networks have historically been used to study information transfer between people, groups, or organizations, they are increasingly employed in a variety of contexts, including the study of co-occurrences of or associations between narrative elements (Smith 2007), cultural objects (Breiger and Puetz 2015), multimedia content (Meng and Shyu 2012), and even food flavors (Ahn et al. 2011) and human genes (Schafer and Strimmer 2005). In the context of music, the nodes in the network would be songs, while the edges between them might represent varying degrees of feature overlap or similarity.

[5] Another familiar metaphor that approximates this idea is that of the cultural "milieu" or "fabric." This concept encompasses the population of cultural products that producers and/or consumers have access to in a given context. In the market for popular music, this might include all current and previously released songs, which can then be connected to one another, however distantly, based on their shared feature sets. While the theoretical and empirical implications of this idea extend beyond the scope of this paper, the imagery of a cultural fabric may help motivate our rationale for extending the concept of networks to cultural products and their constitutive features.

[6] Focusing on songs that appear in the Hot 100 our analysis may suffer from considerable selection bias, an issue we address in Appendix A. However, we believe that any bias in our data does not present a major limitation, as charting songs constitute an appropriate sample for answering our initial research question: what makes popular culture popular? Further, it is consistent with studies that explore the differentiated outcomes for cultural products

that get shortlisted for prizes versus those that win (e.g., Kovács and Sharkey 2014; Sorensen 2007). While other factors such as artist popularity and marketing support play an important role in driving certain songs into the Hot 100, we are primarily interested in understanding why, conditional on entering the charts, certain songs outperform others.

[7] The initial algorithm for determining the charts included a combination of radio airplay and a survey of selected record stores across the country. This methodology had several flaws, as it relied on human reporting for a large portion of the input and was therefore subject to both personal biases and external influence. In November 1991, *Billboard* replaced the self-reported sales data with SoundScan's point-of-sale data from most of the record stores in the United States (for more on the history of the algorithm and the consequences of the shift, see Anand and Peterson [2000]). We run a series of supplementary analyses to test how this development influences our results, and find that the effect of SoundScan falls within the range of expected historical variation across our dataset (see Appendix B).

[8] In case consumer selection decisions occur at the artist rather than song level, we also ran our models using artist-level genre attributions. Results are consistent.

[9] One of the weaknesses in our data is that these genre codes were applied in the early twenty-first century, rather than the year in which each song was originally released. While genre attributions are admittedly dynamic (Lena and Peterson 2008), we believe it is reasonable to assume that historical attributions are for the most part consistent with our data. Furthermore, though genres appear and disappear over the course of our data, and those that persist have evolved, such changes have their provenance predominantly at the sub-genre or "style" level (e.g., "Hard Rock" and "Roots Rock" versus "Rock"). Employing primary, song-level genre assignments means that misattributions are unlikely or should be relegated to fringe cases.

[10] Each of these features is weighted equally to calculate our pairwise cosine similarity measure. While we considered prioritizing certain features over others (e.g., weighting "tempo" more heavily than "mode"), conversations with musicologists and computer scientists specializing in MIR provided no consistent rationale for using weights. Moreover, the eleven features included in our analysis were designed to encompass the most

important dimensions of songs in a relatively evenhanded and comparable way, with the possible exception of mode, which has a slightly outsized influence due to its binary (0,1) rather than continuous scale.

[11] When two songs were the only representatives of their respective genres over the previous year (a rare occurrence, largely confined to the early years of the chart), we used the minimum similarity between any pair of genres from the year prior to the focal song's debut week to construct our weighted measure. For example, if a focal song has a primary genre of "vocal," and is the only such track to appear for an entire year on the charts, then the minimum weight (i.e., the largest distance between two genres' vectors of average features) is used as the weighting multiplier for that song's cosine similarity with every other song on the charts during the previous year.

[12] In another set of models (available by request), we checked the robustness of our results vis-à-vis different levels of reliance on genre classification. To do this, we calculated two additional typicality variables—*all pair typicality (yearly)* and *within-genre typicality (yearly)*—which seek to provide further evidence that our results hold across multiple specifications. *All pair typicality* is again a cosine similarity measure, but it is the simple, unweighted average of each song with all other songs that appeared on the charts in the previous 52 weeks (see also Appendix B). It is our main independent variable without any genre-based weighting. *Within-genre typicality* is, as its name implies, the average cosine similarity between each song and the average feature vector for all other songs affiliated with the same genre in a given year. This version of the measure captures how typical a song is for its given genre. We found substantively similar results using both of these variables across models 3-6.

[13] We are grateful to an anonymous reviewer for bringing this issue to our attention.

[14] All control variables—including genre affiliation, dummies for each musical key (C through B), major label dummy, *long song*, *multiple memberships, crossover, reissue,* and half-decade time dummies—are included in these models, but are not shown in the figure (see Appendix **Table A4** for full results).

[15] In addition to including the crossover dummy in our models, we also ran separate versions of models 4 and 6 for crossover songs and non-crossover songs. Our main findings hold for *non*-crossover songs—they are benefitted by being optimally differentiated—but not for crossovers. However, crossovers do comprise a higher proportion of #1 songs (30%) than their overall chart presence would suggest (24% of all songs are crossovers by our measure).

[16] For this analysis, we exclude 1991 as SoundScan was implemented in November of that year and we wanted to see if there was a discontinuity in our results.

[17] Results are robust to decade- and year-level dummy variables as well.