Ehrenberg-Bass Institute **Working Paper**:

Death by 1,000 'true fans': Do marketing laws apply to music listening?

This working paper, dated 11 October 2022, is forthcoming in the ***Journal of Consumer Behaviour.***

Authors:

Dr Zac Anesbury - Ehrenberg-Bass Institute
Callum Davies - Ehrenberg-Bass Institute
Dr Carl Driesener - Ehrenberg-Bass Institute
Dr Bill Page - Ehrenberg-Bass Institute
Dr Luke Greenacre - - Monash University
Dr Song Yang - Ehrenberg-Bass Institute
Prof Johan Bruwer - Ehrenberg-Bass Institute

University of South Australia
Ehrenberg-Bass
Institute for Marketing Science

Death by 1,000 'true fans': Do marketing laws apply to music listening?

Author Details

1. **Dr Zachary William Anesbury**

Senior Marketing Scientist, Ehrenberg-Bass Institute, Lecturer, UniSA Business

University of South Australia, South Australia, AUS

Level 4 Yungondi Building, 70 North Terrace Adelaide, South Australia, AUS 5000

Zachary.Anesbury@marketingscience.info

2. **Callum Davies**

Marketing Scientist, Ehrenberg-Bass Institute, UniSA Business

University of South Australia, South Australia, AUS

Level 4 Yungondi Building, 70 North Terrace Adelaide, South Australia, AUS 5000

Callum.Davies@marketingscience.info

3. **Dr Carl Driesener**

Senior Marketing Scientist, Ehrenberg-Bass Institute, Senior Research Fellow, UniSA Business

University of South Australia, South Australia, AUS

Level 4 Yungondi Building, 70 North Terrace Adelaide, South Australia, AUS 5000

Carl.Driesener@marketingscience.info

4. **Dr Bill Page**

Senior Marketing Scientist, Ehrenberg-Bass Institute, Senior Research Fellow, UniSA Business

University of South Australia, South Australia, AUS

Level 4 Yungondi Building, 70 North Terrace Adelaide, South Australia, AUS 5000

Bill.Page@marketingscience.info

5. **Dr Luke Greenacre**

Senior Lecturer, Department of Marketing

Monash University, Victoria, AUS

Luke.Greenacre@manash.edu

6. **Dr Song Yang**

Marketing Scientist, Ehrenberg-Bass Institute, Lecturer, UniSA Business

University of South Australia, South Australia, AUS

Level 4 Yungondi Building, 70 North Terrace Adelaide, South Australia, AUS 5000

Song.Yang@marketingscience.info

7. **Professor Johan Bruwer**

Research Professor, UniSA Business

University of South Australia, South Australia, AUS

Level 4 Yungondi Building, 70 North Terrace Adelaide, South Australia, AUS 5000

Johan.Bruwer@unisa.edu.au

Abstract

The advice to musicians and marketers is to focus on what they love: a truism for practitioners is to find 1,000 “true fans” and make $100 from each of them (Kelly, 2008). If this advice is correct, we should see musicians with loyal user bases engaging more with their favourite artists and less with other music, suggesting a narrow targeting strategy would suffice. On the other hand, the established marketing laws indicate that the listeners of very different genres should overlap more than conventional wisdom would suggest, supporting the need for a much broader approach to targeting potential audiences. Given these conflicting views, musicians need to know if they should market to their existing listeners, the listeners of music similar to theirs (i.e., the same genre), or if they should try to reach a much wider audience. We turn to established choice patterns from the marketing literature to address these questions in the music context. This study examines 84,000,000 observations of music listening from 27,000 unique global users between 2013 and 2014 and survey data from 2019 containing music listening from over 1,000 representative respondents in the United States. The results show that listening follows the Duplication of Purchase law for genres, artists, albums, and songs, at an annual, six-months, three-months, one-month, and one-week period, with no indication of partitioned music listening. The implication is that musicians should try to reach *all* potential listeners, regardless of what they already listen to. These findings contribute to the theoretical knowledge about duplication analyses of various durations, extend the contexts of choice behaviour that exhibit this pattern, and managerially, to knowledge about the extent of potential audiences and ‘share of ear’ competition.

Keywords: music, duplication of purchase, partitions, groupings, marketing laws

1 Introduction

Music listening is a popular and prevalent human choice activity. Americans spend about 32 hours a week, or four-and-a-half hours a day, listening to music (Nielsen, 2017). There are over 250 million paid music streaming service subscribers, accounting for 37% of the US$19B industry (International Federation of the Phonographic Industry, 2019). Streaming technologies provide unprecedented access to millions of songs (Krause et al., 2015), making finding music a potentially overwhelming choice. Record companies invest at least US$5.8B annually on new artist scouting and marketing activities (International Federation of the Phonographic Industry, 2019).

Consequently, understanding how music competes for listeners can maximise an artist's growth potential and stakeholder profitability. In the context of the growing popularity of music streaming and the subsequent availability of millions of songs, understanding patterns of music consumption behaviour is more important than ever for musicians (Wlömert and Papies, 2016). Still, a further complication makes this an exciting activity for study: how does an artist get listened to in an age where algorithmic recommendation services shape listening, which may reduce exposure to new music?

A popular strategy recommended by music pundits is that a band needs only 1,000 true fans to succeed: if each of these fans spends $100 a year, the musician generates an income of $100,000 (Kelly, 2008). The strategy is rarely questioned, beyond caveats such as the need to pay processing fees, and overheads like studio space, equipment costs, or multiple band members are "easy" changes: you would "simply" need 2,000 fans for a band with two members. When this concept is set against the backdrop of huge potential audiences and the potential for platforms like Patreon to fund content creators, it is seen as "a much saner destiny to hope for. And you are much more likely to actually arrive there" (Kelly, 2008).

But, for this to be true, music consumption would have to look very different to the consumption of media, consumer packaged goods, durables, store choice, and business-to-business (B2B) contexts (see Table 1), which is, of course, entirely possible. We would expect that if the

1,000 true fans concept is useful, we would see that listeners would show firm favourites: they would be more likely to listen to smaller groups of their favourite artists or albums, with artists and genres insulated from each other in terms of audiences (i.e., there is partitioning). We would not see broad, multi-genre listening. The current paper thus forms an initial face validity test of the 1,000 true fans hypothesis.

To understand how music competes, we turn to the fitted Negative Binomial Distribution (NBD)-Dirichlet and its outputs. The fitted NDB-Dirichlet, a zero-order model of competitive market structure, is known to fit many competitive contexts (Driesener and Rungie, 2022; Driesener et al., 2017). One of the outputs of this model is benchmarks for observed sharing of customers between brands, known as Duplication of Purchase. The Duplication of Purchase law (DoP) is a highly generalised pattern that provides benchmarks for competition, to the extent that it has become one of the very few Laws in marketing. This empirical generalisation helps consumer goods brands understand which other brands their customers are buying (Danaher et al., 2008; Reuver and Jonkheer, 2017), assessing the extent to which brands share customers or are insulated from each other. For this research, DoP may clarify a different choice context: the level of listener sharing in music. In the absence of an appropriate generalised term, in this manuscript, we use 'brand' to indicate the genre type, artist name, album name, or song name, depending on the context. DoP states that the level of customer sharing (or duplication) between two 'brands' is proportional to the size of the second brand's customer base (its penetration) (Goodhardt and Ehrenberg, 1969). This generalisation identifies that more of a brand's customers buy larger brands, and fewer buy smaller brands (Dawes, 2008; Ehrenberg et al., 2004), and this difference is predictable. Brand size dominates the sharing of customers. As DoP provides a benchmark for competition, there is the possibility of benchmarking the "neutrality" of recommendation engines. If an artist pays to be promoted, it could be used to determine if the algorithm segments a market around bands who pay more or not.

However, music listening may not show the expected DoP pattern. Streaming service recommendations, personalised playlists, and album-based listening (Domingues et al., 2013; Ricci et al., 2015; Yapriady and Uitdenbogerd, 2005), let alone personal preferences, may violate the underlying choice patterns from which the DoP pattern arises in other areas of consumer choice. Most people would assume jazz and hip-hop to have two distinct groups of listeners; we could similarly suggest classical and pirate metal might have two distinct listener groups.

While previous research has examined media consumption, for example, television (Ehrenberg and Goodhardt, 1969) or websites (Corkindale et al., 2013), this study provides new knowledge by examining DoP for music genres, artists, albums, and songs, during various durations (six months, three months, one month, and one week). DoP analyses typically use a purchase window of 12 months (Ehrenberg et al., 2004) to capture seasonality and allow for repertoire buying to be expressed in the data by using a period representing multiple average inter-purchase intervals (Banelis et al., 2013; Trinh, 2014). Analysing shorter periods may distort the pattern. By assessing music genre consumption over different periods, this paper extends the known boundary conditions of the DoP pattern to a novel consumer choice context and assesses the impact of varying lengths of time.

2 Background and Research Questions

2.1 Duplication of Purchase

Duplication of Purchase (DoP) does two things. Firstly, it identifies the importance of brand size (i.e., penetration) in determining customer sharing (or customer overlap) between any given pair of brands. The generalisation is that the brand's size is the dominant factor when determining sharing. Secondly, an expected level of customer sharing with competitors derives from the law against which actual sharing can be compared (Anesbury et al., 2021; Ehrenberg et al., 2004; Sjostrom et al., 2014; Wilson and Winchester, 2019). Originating from media audience research (e.g. Agostini, 1961; Chatfield and Goodhardt, 1975; Driesener and Rungie, 2022; Goodhardt, 1966; Naami et al., 2021; Winchester and Lees, 2013), the number of viewers that watch any program pair is predicated on the latter program's audience size (Goodhardt and Ehrenberg, 1969). This pattern also exists when examining television programs and channels (Collins et al., 2003) and radio (Lees and Wright, 2013).

The pattern's successful replication occurs in many situations (see Table 1). These include, but are not limited to, entertainment product categories - gambling in the United States, Australia and Macao (Lam, 2006; Lam and Ozorio, 2013), tourist destinations (Mansfield and Romaniuk, 2003; Mansfield et al., 2003), and high-involvement purchase contexts such as aviation fuel contracts (Ehrenberg, 1975; Uncles and Ehrenberg, 1990b). More recently, its scope for application has been expanded in a novel application examining brand extensions in different categories (Grasby et al., 2022).

When the pattern holds in a market, a key implication for practice is that brands (or, in this instance, bands) compete head-on, in line with size. That is, the biggest competitor for all brands in a market is *always* the biggest brand: no brand is insulated or protected from competition in some way. Knowing that the pattern fits in a market changes the marketing problem from one of finding a "niche" (or your 1,000 true fans) into an all-out battle for listening: competing for Mental and Physical Availability (Romaniuk and Sharp, 2021; Sharp, 2010) in an unsegmented market – that is competing for a wide rather than a narrow audience. This change in thinking shapes the strategies marketers can use to grow market share: understanding if and how Duplication of Purchase (or Listening) fits is crucial for guiding a successful marketing strategy.

---Table 1---

2.2 Deviations from the Duplication of Purchase law

While sharing in line with size is the predominant finding, deviations from this pattern can be found by comparing actual sharing with expected sharing. These deviations are classified as either groupings or partitions. Groupings occur when a sub-set of brands share more customers with one another (perhaps indicating enhanced competition), but the sharing with brands outside the sub-set is as expected (no lessening of competition with other brands). Partitions, although rare (Scriven et al., 2017), occur when (a sub-set of) brands with higher-than-expected sharing share many fewer customers with the remaining market than expected (indicating these brands are somewhat insulated from competition). Partitioning can result from functional differences causing higher substitutability amongst those sharing the similarity (Lees and Wright, 2013; Romaniuk and Dawes, 2005), for example, instant caffeinated vs decaffeinated coffee (Kalwani and Morrison, 1977).

Identified partitions include leaded vs unleaded petrol (Bennett et al., 2000; Scriven and Ehrenberg, 1994), luxury vs non-luxury automobiles (Ehrenberg and Bound, 2000), country of origin, vehicle type, and specific company's automobiles (Colombo et al., 2000), filtered vs non-filtered, and menthol vs non-menthol cigarettes (Carter and Silverman, 2004), geographically close tourist destinations (Mansfield et al., 2003), Chinese vs global television brands (Bennett, 2008), and talk vs music radio stations (Lees and Wright, 2013). So, while partitions, or insulated 'audiences', are rare, they can occur.

The assumption that jazz and hip-hop would have two distinct groups of listeners or that classical music listeners would not listen to pirate metal suggests that partitions would be present in the music context.

2.3 Why DoP might not apply to music listening

2.3.1 Music choice may not follow a zero-order process

Many factors may mean the zero-order processes underlying choice in other behavioural contexts may not be present in music listening. First, many people consume music online via streaming services, which combat the overwhelming choice between countless songs by modelling user preferences to suggest music (Domingues et al., 2013). Recommendation systems include collaborative filtering (i.e., groups users on taste similarity, see Yapriady and Uitdenbogerd, 2005), content-based filtering (i.e., music stylistically like previously enjoyed, see Ricci et al., 2015), and other approaches (e.g., demographic profiling, knowledge-based, constraint-based, community-based, and hybrid systems, see Ricci et al., 2015). Second, common features, including Spotify's "Discover Weekly" (i.e., individually personalised playlists and auto-play), attempt to recommend new music to the user weekly. Third, when listening to complete albums, the following songs are from the same artist. Finally, people are likely to listen to similar music in a thematic or auditory sense: listening to one song may make them more likely to listen to another similar song and less likely to listen to a less similar song. Thus, streaming service recommendations, personalised playlists, album listening, and the nature of musical preference may violate the Dirichlet assumption of as-if-random (zero-order) brand (music) choice. Consequently, DoP may not hold for music listening.

2.3.2 Music – one category, many sub-categories

While manufacturers of "normal" consumer products, that is, consumer goods brands, may claim some familial connection between their products through master-brand strategies, brand and line extensions, or merely through variants of the same brand, music is potentially different. The first potential difference is the issue of provenance: songs made by a single artist are likely to sound similar because an individual or small group of individuals usually use the same instruments, voices, themes, and styles. The second is that songs are traditionally also bundled together as an album, and even when selling one song (i.e., a "single"), more than one song is present on the media (e.g., compact disc or cassette tape). In a streaming sense, the song selection interface design often encourages listening to the other songs on the album or by the same artist and playlists featuring the "best of" a particular artist encourages further listening at that same level. At the artist or album level, we would expect to see higher than usual partitioning of listening if there is no album 'bundling' or other connections between songs. If all artists need is the support of 1,000 true fans, we would need to see some evidence of partitioning, at least at the artist level.

At the genre level, we might also expect that genres operate as distinct sub-categories as the musical styles of different genres vary enormously. For artists attempting to grow, it would be helpful to know if they should promote their music to people who are already listeners of similar music, or should they also search more broadly for new listeners.

We, therefore, look at 'duplication of listening' at the most atomic level (songs) and an album, artist, and ultimately the most macro-level: genre.

RQ1: Does annual music listening follow the law of Duplication of Purchase for (a) genre, (b) artist, (c) album, and (d) songs?

2.3.3 DoP might not apply for various durations

DoP analyses are typically for 12-month periods (Brewis-Levie and Harris, 2000; Lam and Ozorio, 2013; Mansfield et al., 2003), but shorter durations are critical for this context. Penetration, the percentage of people that listen to a music 'brand' at least once, is, after all, dependent on the duration (Ehrenberg et al., 2004). As time increases, the penetration of each music genre, artist, album, or song builds from more occasional listeners. For consumer goods, longer durations relate to increased individual brand penetrations as household repertoires expand (Ehrenberg, 1988), the increased number of unique brands bought (Banelis et al., 2013; Trinh, 2014), and increased customer-sharing between brands as buyers switch between more alternatives over extended purchase sequences (Stern and Hammond, 2004).

Duration also influences the Duplication Coefficient. This competitive intensity measure is the division of the average observed duplications by the average penetration of all brands (Dawes and Nenycz-Thiel, 2013; Ehrenberg and Bound, 1999; Ehrenberg and Pouilleau, 1992). In a consumer goods context, for an annual period, the average Duplication Coefficient is 1.4 (Ehrenberg, 1991; Keng et al., 1998; Scriven and Danenberg, 2010; Uncles and Ellis, 1989; Wright et al., 1998). When the duration halves, the average decreases to 1.1 (Keng and Ehrenberg, 1984; Wrigley and Dunn, 1984a; Wrigley and Dunn, 1984b). While perhaps not entirely like-for-like, consumer goods research provides valuable benchmarks for the various durations.

In a music listening context, in one week, there is a small opportunity for a listener to listen to any particular song (i.e., lower penetrations), and there is also less opportunity for a listener to listen to additional music (i.e., lower duplications). The ability for further listening beyond the first increases as duration increases; listeners can listen to more additional genres, artists, albums, and songs with an increased period. Therefore, after an initially high Duplication Coefficient (high penetrations, low duplications) at shorter periods, this could plateau with increased periods allowing listeners to share their listening more widely. Therefore, our second research question is:

RQ2: Does music listening follow the law of Duplication of Purchase for (a) six months, (b) three months, (c) one month, and (d) one week?

2.3.4 Music partitioning

Suppose Kelly's (2008) 1,000 true fans concept were to be useful. In that case, we should see this heavy-user loyalty creating partitions (or unique audiences) in the data around particular artists: smaller 'niche' artists or albums, as well as blockbusters. If we find clear partitioning, artists and those marketing them should focus their efforts on their existing listeners, exploiting the '1,000 true fans'. Suppose it does not exist, and competition occurs in line with penetration. In that case, music marketers should focus on a broad reach strategy, where new listeners are found among existing listeners of *all* music, rather than merely those of the current artist or genre.

We might also see partitioning between genres. Should music marketers seek to reach listeners who already listen to music of a similar genre, or could anyone who listens to music be a target? If individuals are fans of a single genre, that is, they listen to a particular genre more than others, this would necessitate that they decrease their listening of other genres, and so we could see this in the data. If strong partitioning exists between genres, musicians and those marketing them should target their marketing efforts on people who already listen to similar genres of music and avoid wasting spending on people who listen to other, safely partitioned genres. But, if there are no partitions, music marketers should focus on a reach strategy.

Individuals may also listen to stylistically similar music (e.g., Blues listeners may be Jazz listeners). Previous research used factor analyses to load genres onto unified style dimensions – Grunge, Heavy Metal, Punk, Alternative and Classic Rock onto a "rebellious" dimension (George et al., 2007), and Blues, Jazz, Classical and Folk onto a "reflective and complex" dimension

(Rentfrow and Gosling, 2003). Loadings indicate functionally similar genres where individuals listen to music in that dimension and potentially may avoid others (i.e., partitioning).

Finding clear partitioning in this context (a failure of the DoP law) would mean we have identified a boundary condition for DoP analysis in this choice context. Further interesting questions would then stem from this: is the failure of DoP in this instance a function of music choice being different from other consumer choice behaviours, and/or are differences amplified by the recommendation engines used by streaming services disrupting the as-if random and zero-order assumptions underpinning the use of DoP? The absence of these partitions in a world of such recommendation engines potentially calls into question their value or perhaps supports the primacy of natural human tendencies to seek a variety (e.g. Kahn, 1995) and avoid mundane repetition. Consumers may not want a bland diet of sameness served up to them, regardless of what marketers and technologists seem to think or try to achieve.

With regards to music consumption, if we find a broad, unpartitioned market, the implication would be that people must have repertoires of "brands" (i.e., artists, songs, and genres) that they listen to, but similar to brand buying, they have propensities to behave outside of their normal behaviour, and "buy" (i.e., listen) outside of their normal repertoire. The behaviour change might be due to different moods, suggestions from friends, or following the whim of an algorithm rather than a brand being out of stock or on promotion. Our final DoP research question is:

RQ3: Is there evidence of partitioning between genres?

3 Method

3.1 Data

3.1.1 Data Set 1: observational music listening data

Consistent with the multiple sets of data approach (Barwise, 1995; Lindsay and Ehrenberg, 1993), the study examines four individual data sets collected between 2013 and 2014. The data were obtained from Last.fm using their publicly accessible Application Programming Interface (API). Last.fm is an online personalised radio station and music recommendation service that collects individual-level music listening data worldwide. Users of Last.fm agree to have their listening behaviour tracked to receive recommendations, share music, and for their personal interest in their listening behaviour. Last.fm supports numerous streaming services, including Spotify, YouTube, Apple Music, Google Play Music and Pandora (Last.fm Ltd, 2019).

The constraints of the API meant that there are limitations to harvesting the listening histories of members of the website *en masse*, but rather individual identifiers must be found through "group" membership and then listening histories. Therefore, our first data represent the listening of self-selected members of two of the largest genre-based "interest" groups: 'Indie & Alternative' and 'Heavy Metal'. We also selected two further groups we thought would have the highest chance of being different individuals with very different preferences: 'Classical' and 'Trance'. However, it is essential to note that the group 'Heavy Metal' is a discussion group about heavy metal, one of many that each individual may have joined. Groups not related to genre also exist, such as groups to discuss upcoming events in particular regions. The data contains each listener's history of music consumption across all songs, albums, artists, and genres. We retain the name as a convenient handle. An analogy would be an animal ownership database where people could call themselves a 'cat' person, but the database captures all pets, such as dogs, fish, birds, and rodents, as well as cats.

Seeking to maximise the chances for partitioning, we merged all four groups into a single data set. The creation of this fifth dataset, referred to as "Last.fm," gives the maximum possible chance for partitioning. In total, the data captured approximately 84 million listens from about 27,000 unique users over one year. "Listens" are recorded when users play a track for longer than 30 seconds, play for at least half of a track's duration, or play for at least 4 minutes (Last.fm, 2013).

Although this data allows for the examination of actual listening behaviour, the nature of the data (i.e., self-selected, people choose to join the group and can join multiple groups, and the initial choice to join Last.fm), there may be representativeness issues in terms of both weight of listening and demographic skews. The requirement for individuals to know about Last.fm and install the app makes it likely that they are more engaged with music and may listen to more music than the general population. We also cannot determine the representativeness of the data regarding age, gender, or any other demographic variable.

3.1.2 Data Set 6 survey music listening data

Due to the potential for the Last.fm data to be weighted towards heavy music listeners and the lack of information regarding representativeness at a population level, we also conducted a separate data collection phase to obtain a sixth data set. We recruited participants via Toluna, a reputable commercial panel provider that works with corporate clients providing global data collection and research capacity. All respondents were 18+ and had listened to any music in the past week. We collected genre listening in the last week, and respondents also answered demographic questions about age, gender, residential state, and social class. The final data (n=1,036) consists of a nationally representative US sample by age, gender, race, and location – see Table 2. The age ranges from 18 to 84, mean age = 48.6, SD = 17.6 (U.S. Census Bureau Population Division, 2019).

---Table 2---

3.1.3 Genre Classification

Genres are common in examining relationships between music preference and demographic and psychographic variables (e.g. Peterson and Kern, 1996; Rentfrow and Gosling, 2003; Savage, 2006). With many brands[1] at the genre, artist, album, and song level, even with the extensive data set, many have small listening bases leading to small sample errors. Artists and listeners are free to use their own genre classifications, which can quickly become quite specific: the genre "metal" splinters in the data into industrial metal, metalcore, alternative metal, nu metal, symphonic metal, death metal, melodic death metal, gothic metal, heavy metal, progressive metal, deathcore, thrash metal, power metal, doom metal, black metal, pirate metal, Viking metal, technical death metal, groove metal, folk metal, and love metal, among others. However, rather than understanding how

[1] We use 'brand' to mean the genre type, artist name, album name, or song name, depending on the context.

all possible genres share listeners, each (sub)genre was recoded into a "meta-genre", reducing the extensive list to 13.

Prior studies typically measure different genres and use different classifications. For example, Savage (2006) uses nine genres, none of which are Pop, while North (2007) uses 19 genres, including four variations of Pop. Rentfrow and Gosling (2003) generated and tested a list of genres to use in their Short Test of Musical Preferences (STOMP), consisting of 14 genres that load onto four music dimensions. These dimensions are (1) reflective and complex (Classical, Jazz, Blues, and Folk), (2) intense and rebellious (Alternative, Rock, and Heavy Metal), (3) upbeat and conventional (Country, Pop, Religious, and Soundtracks), and (4) energetic and rhythmic (Rap/Hip-Hop, Soul/Funk, and Electronica/Dance). This current study uses classifications similar to STOMP (Rentfrow and Gosling, 2003; Rentfrow and Gosling, 2006) but also uses genre names more closely aligned with Apple Music and Spotify to be as familiar as possible to respondents and provide some degree of comparability. Our classifications allow the grouping of the many self-classified genres into meta genres (e.g., pirate metal, Viking metal, and at least a dozen more into Metal) and shape the questionnaire for our data collection. Within the DoP literature, using predefined categories, for example, 'bread and bagels' from Nielsen Kilts Centre for Marketing (Grasby et al., 2022) or 'fashion clothing' from Kantar World Panel UK (Chowdhury et al., 2021), is the norm.

Our final list of 13 genres derived from STOMP (Rentfrow and Gosling, 2003) and the classifications of Apple Music and Spotify are Indie/Alternative, Blues, Classical, Electronica/Dance, Folk, Metal, Hip-Hop/Rap, Jazz, Pop, Rock, R&B/Soul/Funk, and Soundtracks. Recoding (sub)genres into meta-genres relied on the STOMP/commercial genre list and used consensus among the research team. We also note that as a sub-genre became more specific or harder to classify into our genre schema, it represented fewer tracks and had less of an effect on the overall listening patterns.

3.2 Analysis Method

This research uses duplication of purchase analyses (Ehrenberg, 1988; Goodhardt and Ehrenberg, 1969). For DoP. We examine the five sets of Last.fm data (i.e., all listening data from all Last.fm users) using various periods (i.e., 2x6 months and 4x3 months plus one randomly selected 1-month and 1-week slice). The sixth data set only had listening information for the 'last week'.

3.2.1 Analysing the Duplication of Purchase (DoP)

To demonstrate how bands compete, we apply DoP (Equation 1) to claimed listening behaviour of genres from the survey administered online in the United States (i.e., data set 6).

Equation 1 - duplication of purchase

$$b_{x/y} = Db_x$$

where

$b_{x/y}$ is the proportion of brand y's customers that also buy brand x
b_x is the proportion of the sample who buy brand x at least once in a given period
D is the duplication coefficient which is the average observed duplications between each pair of brands divided by the overall average penetration (Ehrenberg and Pouilleau, 1992; Lees and Wright, 2009)

Table 3 shows how all genres share, on average, 63% of listeners with the largest genre, Rock, 57% with Pop and so on. Overall, the smallest genres, Folk and Electronica/Dance, share larger listener proportions with Rock and Pop, the largest genres. In contrast, Rock and Pop share fewer listeners with Folk and Electronica/Dance. Overall, sharing occurs in line with size.

---Table 3---

3.2.2 Tests of fit: Does 'brand' size drive listener sharing?

We undertook several tests to determine whether the observed data follows a DoP pattern. We first calculated the correlation between average duplication values (i.e., the average of the columns in Table 3) and brand penetrations. We then calculated the duplication coefficient (D) by dividing the average of the observed duplications by the average penetration of all brands. The D-coefficient measures the level of competitive intensity between brands (Dawes and Nenycz-Thiel, 2013). For example, a D-coefficient of 2 means that the proportion of Brand 1's listeners listening to Brand 2 will be double the penetration of Brand 2. Finally, the D-coefficient is multiplied by each brand's observed penetration, providing the expected duplication (see Equation 1). Differences between the observed and expected duplications are reported as mean absolute deviations (MADs) and mean absolute percentage errors (MAPEs). As "all other" is an aggregation of numerous brands and not an actual brand, we exclude it from tests of fit measures.

There is no consistent approach to assessing deviations in the literature - some adopt a 10-pp threshold (e.g., Cohen et al., 2012) while others assess the general pattern (e.g., Dawes, 2008). This study adopts a 10-pp threshold for managerially significant deviations as it is quantifiable and tractable given the numerous datasets and analyses executed here.

3.2.3 Testing if 'brands' partition from the rest of the market

Deviations occur when brands share more or fewer customers than expected. Such deviations from the pattern can indicate the presence of submarkets or partitions (Sharp and Sharp, 1997) where 2+ brands share more customers than expected but share normally with all others. Typically, partitions occur when 2+ brands share more customers but share fewer with the rest of the market[2]. Partitioning is often attributable to functional differences when customers desire specific attributes offered by a brand or a subset of the brands (e.g., talk-back and music stations Lees and Wright, 2013). Knowledge of the presence of partitions is useful to managers because they can expect to gain fewer customers from these brands or, if their brand is part of a partition, be insulated from external competition. Their presence or absence is the key to our analysis of music choice behaviour and thus understanding it. The Partition Sharing Index (PSI) determines if partition exists (Anesbury et al., 2018; Anesbury et al., 2021; Naami et al., 2021; Sjostrom et al., 2014). PSI shows the level of sharing between two brands, given their size. Equation 2 presents the calculation:

[2] It is also possible for a single brand to partition from the competition.

Equation 2 – Partition Sharing Index

$$PSI_{ij} = \frac{S_{ij}}{(D \times p_i)}$$

Where S_{ij} is the observed duplication between brand i and brand j, D is the duplication coefficient (Ehrenberg et al., 2004), and p_i is the penetration of brand i (Sjostrom et al., 2014).

The averaging of the PSI calculates the level of sharing occurring within a group of brands (intra-partition) and the level of sharing between two groups (inter-partition). The index shows whether two groups share more or fewer customers than expected (Sjostrom et al., 2014).

A PSI of 1.0 means the group of brands share as expected. A PSI of 1.5 means the brands share 50% more listeners than expected, while a PSI of 0.5 is 50% less than expected. Consistent with previous studies, brands sharing 20% more or fewer listeners (PSI ≤ 0.8 or ≥ 1.2) are considered managerially significant (Anesbury et al., 2018; Sjostrom et al., 2014). Partitions exist when the brands have an intra-PSI of 1.2 or higher *and* an inter-PSI of 0.8 or lower, where prior research consistently defines these as being managerially significant (e.g., worthy of further investigation and unlikely to be caused by sample issues) (Anesbury et al., 2018; Anesbury et al., 2020; Sjostrom et al., 2014). Groupings occur when the brands have an intra-PSI of 1.2 or higher, *but* the inter-PSI is greater than 0.8.

Intra-PSIs were computed for each sub-genre grouping to indicate the degree to which sub-genres belong to the same meta genre and share listeners. We also calculate the inter-PSIs for each sub-genre grouping to determine the level of sharing occurring with sub-genres outside that grouping (i.e., the "rest of the market"). In addition to calculating inter-PSIs between one sub-genre grouping and the rest of the market, we calculate inter-PSIs between individual groupings.

Genres that belong to the same grouping should hypothetically overshare listeners with each other and "undershare" with genres outside of this grouping. Therefore, using Rentfrow and Gosling's (2003) four music preference dimensions, the following proposed oversharing genres also tested are (1) Reflective & Complex: Classical, Jazz, Blues and Folk, (2) Intense & Rebellious:

Indie/Alternative, Rock and Heavy Metal, (3) Upbeat & Conventional: Country, Pop, Religious and Soundtracks, and (4) Energetic & Rhythmic: Hip-Hop/Rap, R&B/Soul/Funk, Electronica/Dance. On the other hand, the following genres should "undershare" listeners with each other based on Savage's (2006) findings: (1) Classical with Rock and/or Electronic/Dance, and (2) Country with any other genre, excluding Jazz.

4 Results

4.1 Music listening follows the general DoP pattern

Broadly, we demonstrate the pattern predicted by the law of duplication of purchase (DoP) by much higher listener sharing with the largest brand (62% on average) than with the smallest brand (5% on average). Specific tests of fit can assess the applicability of DoP to music listening (see Table 4). Firstly, Pearson's correlations between average penetration and average sharing assess the overall match between brand size and observed sharing. These correlations are both consistently high and positive, with the overall average Pearson's correlation across 21 datasets ranging from 0.94 to 0.99, showing the consistently strong relationship between brand penetration and average listener sharing.

Secondly, theoretical average duplication values estimated by the DoP formula should be close to the observed values. The test compares the mean absolute deviations (MADs) and mean absolute percentage errors (MAPEs). MAD is the average amount that a value deviates from the overall mean (Gorard, 2005), while MAPE is the amount that deviates from the overall mean, expressed as a percentage (Armstrong and Collopy, 1992). These measures should be low if the observed and expected sharing levels are similar, which indicates a good fit.

Table 4 shows the observed and expected sharing display a close fit, as indicated by the MADs, ranging from 1-5 (average MAD = 3). MAPEs between observed and expected are generally low (average MAPE = 10). However, instances, where MAPE is high, may indicate a departure from the pattern.

---Table 4---

We first turn to *RQ1 – does music listening follow the Law of Duplication of Purchase?* Compared with consumer goods, the D-Coefficient obtained from music listening behaviour is high. 18 of 21 D-Coefficients exceed that of the average consumer packaged goods D-Coefficient for 1 year. Furthermore, average duplications are quite high, indicating considerable sharing between all genres, artists, albums, and songs. We also note that the more granular analysis (songs) has lower average penetrations and higher D-Coefficients than the more aggregated genre. In conjunction with the high correlations between penetration and sharing, it suggests again that there is considerable sharing between songs. While this may not be a surprising finding for DoP researchers; however, for those in the music industry, the finding likely is.

The MAPEs are higher for albums and songs in the Classical and Trance genres. However, a known MAPE weakness is that it is more sensitive (i.e., produces higher values) when expected values exceed the observed values than when the expected values underestimate observed values (Armstrong and Collopy, 1992). Given that penetration is critical in estimating the expected sharing, MAPEs are higher for lower penetration brands (Wright et al., 2002), impacting the overall fit. Including brands above 10% penetration reduces the MAPE. However, Pearson's correlations are slightly lower, especially for songs (0.97 on average), as fewer brands reach ≥10% penetration criteria and, therefore, a smaller sample of brands. Thus, we find support for RQ1 - music listening follows the law of Duplication of Purchase for (a) genre, (b) artist, (c) album, and (d) songs.

To answer *RQ2 - does music listening also follow the law of Duplication of Purchase for (a) six months, (b) three months, (c) one month, and (d) one week* we analyse listening at the genre, artist, album, and song level (Table 5). The overall average MAD is 3, the MAPE is 13%, and the Pearson's correlation is 0.88, demonstrating a good fit. There is a large degree of similarity in the D-Coefficients across periods, indicating stability in sharing and music listening. And again, the D-Coefficients are generally higher than the identified consumer goods benchmarks.

---Table 5---

Finally, to answer RQ3 - is there evidence of any partitioning between genres, we use the intra- and inter-partition sharing indices (PSI) to determine whether any submarkets or partitions are present. Intra-PSI shows how sub-genres belonging to the same meta-genre share listeners (e.g., how Soft Rock and Classic Rock share listeners under the Rock meta-genre). Inter-PSI shows how this group of sub-genres share listeners with any other genre (e.g., Rock sub-genres with all other genres in the market).

Table 6 shows that sub-genres overshare listeners with sub-genres belonging to the same meta-genre (i.e., Soft Rock with Classic Rock); there is no sub-genre partitioning. For example, Blues sub-genres share, on average, 60% more listeners with each other (1.6 intra-PSI), but the inter-PSI of 1.1 indicates that Blues sub-genres share 10% more listeners than expected with other sub-genres. On average, sub-genres that are part of the same meta-genre grouping share 20% more listeners than expected with each other (intra-PSI = 1.2) but share with any genre outside of the grouping directly proportional to their penetration (inter-PSI = 1.0). Therefore, no meta-genre is insulated – partitions generally do not exist at the genre level.

---Table 6---

5 Discussion

This research shows that the sharing of music listeners between genres, artists, albums, and songs is largely predictable. The findings are essential to music stakeholders concerned with promoting music brands (e.g., artists, music marketers/labels, streaming services). DoP provides an initial assessment of the choice behaviours in music listening, thus potential pathways for brand growth and an understanding of market norms.

5.1 Theoretical Contributions

Our research describes music listening behaviour and examines competition between offerings in genre, artists, albums, and songs. Our results support earlier findings regarding the applicability of the DoP law in media (Goodhardt and Ehrenberg, 1969; Lees and Wright, 2013), entertainment (Lam and Ozorio, 2013), tourism (Dawes et al., 2009) and consumer packaged goods contexts (Bass, 1974). And the findings are consistent with radio station listening research (Lees and Wright, 2013).

This research expands the applications of the Law of Duplication of Purchase beyond consumer goods purchasing and traditional media consumption contexts into a context with thousands (if not millions) of available choices. While streaming services recommendations, personalised playlists, and album listening potentially violate the Dirichlet assumption of as-if-random (zero-order) music (brand) choice, DoP, as a benchmark for the expected competitive structure, remains applicable in this novel extension. This means that the algorithms that attempt to influence choice either perform in ways that represent normal consumer choice. If they are designed to promote specific music genres or artists to create user segments or partitions of heavier listening, then perhaps they aren't working.

In addition, although more music and a greater variety of music is listened to as the duration increases (as is the case with brand buying (Banelis et al., 2013; Trinh, 2014)), the DoP pattern was still as evident in one week as it was in one year.

While Lees and Wright (2013) found evidence of partitioning by format between talk and music stations, we find no evidence of partitioning music genres within our study. While sub-genres overshare listeners with another sub-genre belonging to the same meta-genre (e.g., Soft Rock and J-Rock grouped as Rock), they are not partitioned from the market and share as expected with other genres. Even Classical and Electronica/Dance do not under-share listeners. The pattern is that duplication of listening of genres, artists, albums, or songs, is primarily determined by the overall listener base size. These current findings generate new empirical evidence to verify the generalisation of the television audience research (Ehrenberg and Goodhardt, 1969; Goodhardt, 1966; Goodhardt and Ehrenberg, 1969) to music listening. However, the average MAD is 3, compared to 1. A potential cause may be the many music 'brands' and, therefore, greater sharing variability. Nonetheless, people apparently do not limit themselves to a particular music style. They listen across the whole spectrum of music. Classical listeners can and do listen to Death Metal – in line with the size of the genre and how much overall listening they do. Our results suggest that listeners play Taylor Swift, Beyonce, and Rihanna between Mozart, Beethoven, and Puccini.

5.2 Practical Contributions

The size of the audience drives competition between genres, artists, albums, and songs in the observed period, and this finding provides benchmarks for growth and an understanding of where sales (or streams) are likely to be gained or lost (Dawes et al., 2009). Therefore, artists should expect to share most of their listeners with larger artists and far fewer with smaller ones. That focus should shift away from similarly positioned artists and towards competition with the large music 'brands' (Dawes, 2008). Further, since people's choice of music is predicted and driven by popularity, targeting efforts in so-called differentiated marketing by music may not be effective.

The finding that penetration drives competition in music provides realistic growth strategies. With most listeners listening to the large music 'brands', future listeners are more likely to come from all category listeners, not repeat listeners (Bennett, 2008), and sole listening is not prevalent. Given that listening is not partitioned but grouped in aggregate, an already "big" artist attempting to grow needs to compete for listeners and will do so by taking them from the smaller 'brands'. If an artist is "small", then they *also* need to grow their listener base by over-indexing on listeners; they need to grow by increasing the number of listeners by appealing to the mass market of all possible listeners. Focusing on existing listeners or adjacent genres/styles is a pathway of limited growth compared with competing for listeners in a larger audience.

Our results contrast with Kelly's (2008) '1,000 true fans' theory which posits that sustainable living via 1,000 hardcore fans each spending $100 a year on products (e.g., songs/albums, concerts, merchandise) is possible. A thousand "true" fans might seem feasible, but it will be far more difficult than acquiring more infrequent listeners. We suggest that rather than trying to extract $100 from 1,000 listeners, artists need to understand that it is not possible to achieve an exclusively heavy fanbase: in order to have 1,000 very heavy listeners, they will need to have more medium and a great deal more light listeners. One could consider how many fans overall an artist would need to get to 1,000 true fans (i.e. heavy buyers): we suspect that it would echo the findings for brand buying, and artists would need a substantial proportion of light and ultra-light listeners to reach their 1,000 "true fans" (Dawes et al., 2021).

From the perspective of a music streaming platform (e.g., Spotify and Apple Music), it is crucial to consider that the listeners of one genre are not exclusively listeners of that genre. In our results, over half of all music streamers listen to Rock at least once a year; they also listen to other genres – predictably so. Over half of them will listen to Pop, and one in five will listen to the Blues; levels of sharing are directly in line with the relative sizes of these genres. Therefore, when new Blues music is released, it would be wise to promote it to recent Blues listeners (about one in six) and listeners of other genres. The same is true for artists, albums, and songs.

6 Limitations and Future Research Priorities

There are three key limitations to this research. First, our data consists of respondents who actively chose to track their music listening, leading to self-selection bias, likely skewing toward heavy users (music listeners) (Meyer-Waarden and Benavent, 2009). Last.fm users might be heavy music category users who are widely known to 'buy' more 'brands' (i.e., listen to a broader variety of music) than light users (e.g. Ehrenberg et al., 1990; Fader and Schmittlein, 1993; Sharp et al., 2002). While the survey data somewhat addresses this limitation, an explicit limitation is that it is claimed behaviour. Future research should first determine if these patterns hold with a representative population of music listeners and their observed behaviours.

Second is the measurement of genres for assessing music preferences. As noted, a single genre cannot encapsulate all artists or musical works (Rentfrow et al., 2011), and there is the assumption that research participants are sufficiently knowledgeable about different genres to indicate their music preferences accurately. Although the study uses previously accepted genres (Delsing et al., 2008; Dunn et al., 2012; Rentfrow and Gosling, 2003), the listener sharing of different artists who produce music in different genres requires future research.

The third limitation is the process of defining and grouping genres. While many sub-genres appear to be self-evident, containing terms that make it clear that Pirate Metal and Viking Metal are both parts of the Metal genre, this is not always the case. Defining a market is not inconsequential or uncontroversial but frequently occurs for items people buy (Chowdhury et al., 2021; Grasby et al., 2022) and music people listen to (Rentfrow and Gosling, 2003; Rentfrow and Gosling, 2006). However, as a genre becomes increasingly obscure, it is also listened to by a smaller proportion of the sample, so its precise location in a classification framework becomes less of an issue. Our research shows that it isn't the positioning of the genre but the audience size it attracts that matters in commercial terms, even though this may be an unexpected message for those in the industry to hear.

Fully understanding the battle for listening referred to earlier is fundamental to providing concrete advice to musicians and those marketing them. This research found that, generally, music competes for listeners in line with size at the song, album, artist, and genre levels. That is, bigger songs, albums, artists, and genres also receive more listens. Future research should test this using alternative methods and apply the law of Double Jeopardy (McPhee, 1963) to music listening, an empirical generalisation beyond the scope of this paper. Similar to the law of DoP, Double Jeopardy is a pattern related to the overall size of a brand (in market share or penetration) where smaller brands not only have fewer buyers but are also bought slightly less often by their buyers than larger brands (Ehrenberg et al., 1990). Therefore, brands vary greatly in the number of people who buy them and only slightly in behavioural loyalty (e.g., repeat purchases). Any variation in loyalty is negligible and predictably varies in line with brand size. Suppose Double Jeopardy holds in a music listening context. In that case, this further supports that the path to brand growth is by increasing an artist's band's penetration, which will result in an incremental increase in behavioural loyalty. It would also provide further disconfirmation of Kelly's (2008) 1,000 true fans theory by diminishing the importance of behavioural loyalty from your fans and placing further importance on penetration.

The Pareto principle should also be applied to music listening in future research. Early application of the principle was that the top 20% of heavy buyers of a brand contributed to 80% of its sales (Dubinsky and Hansen, 1982). However, empirical findings demonstrate that the sales contribution of the heaviest top 20% of buyers is closer to 50% (Anschuetz, 2002; Sharp, 2010; Sharp and Romaniuk, 2016). Pending future research, the effective application of the Pareto principle to music brands may further disprove Kelly's 1,000 true fans theory (2008), such that a brand's most loyal customers are unlikely to be a meaningful source of sustainable income, as they only contribute to around half of its sales. Completing NBD-Dirichlet, Double Jeopardy and Pareto principle analyses will round out the knowledge of listening behaviour and allow music marketers to use existing knowledge from marketing to build more effective marketing plans.

A final limitation of this paper is that it provides no *artistic* guidance to artists. While we know that attempting to appeal to a small group of fans is unlikely to lead to becoming a large and successful "brand", how to go about doing this remains an artistic problem.

TABLES

Table 1: Summary of relevant Duplication of Purchase studies

Author/Year	Consumption Context	Country/Region	Notable *Deviations*
Agostini (1961) Agostini (1962)	Magazines	GBR & FRA	
Goodhardt (1966) Ehrenberg (1966) Goodhardt and Ehrenberg (1969) Ehrenberg and Goodhardt (1969) Goodhardt et al. (1975) Barwise and Ehrenberg (1988) Collins et al. (2003)	Television program and channel viewing	GBR & USA	Higher sharing for daytime programs on weekdays, pairs of programs on the same channel on different days and pairs of programs on the same channel on the same day
Lees and Wright (2013)	Radio stations	NZL	Talk and music radio stations format partitioning
Mansfield et al. (2003) Dawes et al. (2009)	Tourist destinations	USA, EURO & ASIA	Location-based partitioning
Sørensen et al. (2012) Lam and Ozorio (2013) Hand and Singh (2014)	Gambling	GBR, USA, AUS, & MAC	Grouping between horse racing and dog racing, which overshares buyers and under shares with scratch cards, lottery tickets and bingo. Higher social class is associated with higher duplication between racing gambling
Ehrenberg (1975) Uncles and Ehrenberg (1990b)	Aviation Fuel Contracts	EURO	
Colombo et al. (2000) Bennett and Graham (2010)	Automobiles	FRA, GBR, & THA	Clusters of French, German, Japanese and General Motor-made cars Luxury vs non-luxury car submarket
Bennett (2004)	Television Sets	CHN	Higher sharing within the cluster of global television brands
Ehrenberg and Goodhardt (1970) Ehrenberg (1988) Keng and Ehrenberg (1984) Uncles and Ehrenberg (1990a) Keng et al. (1998) Anesbury et al. (2018)	Consumer packaged goods	GBR, USA, & AUS	Price related submarkets
Bennett and Ehrenberg (2001) Bennett and Ehrenberg (2002)	Quick Service Outlets/Fast food services	USA & AUS	
Dawes (2008)	Beer	AUS	Partitioning between beer brands part of the Foster's Group
Cohen and Tataru (2011) Romaniuk and Dawes (2005)	Bottled wine and wine retail stores	FRA & AUS	Partitioning by wine variety and adjacent wine price tiers
Uncles and Kwok (2009)	Retail stores	CHN	

Table 2: Sample representativeness in comparison to US census data.

	Sample (%)	**US Census (%)**		**Sample (%)**	**US Census (%)**
Age			**Gender**		
18-24	13	9	Male	47	49
25-39	26	21	Female	53	51
40-54	26	19			
55+	38	29	**Race**		
			Caucasian or white	80	77
Region			African American or Black	11	13
Northeast	18	17	American Indian or Alaskan Native	2	1
Midwest	21	21	Asian	4	6
South	38	38	Hispanic/Latino/Spanish	6	18
West	23	24	Native Hawaiian or Pacific Islander	<1	<1

Source: (U.S. Census Bureau, 2018; U.S. Census Bureau Population Division, 2017; U.S. Census Bureau Population Division, 2019)

Table 3: Percentage of listener sharing amongst music genres (United States)

Genre	Pen (%)	A	B	C	D	E	F	G	H	I	J	K	L	M
A: Rock	52		55	42	32	28	23	30	20	25	18	20	17	13
B: Pop	46	62		43	42	35	23	21	25	24	18	17	20	13
C: Country	39	55	50		33	26	22	21	19	18	13	17	11	13
D: Hip-Hop/Rap	31	53	63	41		55	22	26	25	24	20	20	25	9
E: R&B/Soul/Funk	28	52	57	37	61		20	22	22	25	28	29	21	15
F: Classical	22	54	49	39	31	25		26	33	26	32	25	20	21
G: Metal	18	85	54	44	45	34	32		28	39	21	26	30	16
H: Soundtracks	18	57	64	40	42	34	40	28		28	26	23	30	20
I: Indie/Alternative	17	75	64	40	44	41	33	41	30		24	30	31	27
J: Jazz	16	60	51	32	38	48	45	23	30	25		52	21	23
K: Blues	16	67	51	42	39	52	35	30	27	33	54		23	29
L: Electronica/Dance	14	66	69	32	58	44	33	40	41	40	25	26		14
M: Folk	10	68	58	49	26	41	45	28	35	45	37	44	19	
Average Duplication		63	57	40	41	38	31	28	28	29	26	28	22	18

Table 4: DoP tests of fit measures for 1-year music listening

Analysis Level	Group name	Average Penetration (%)	Average Observed Duplication (%)	D-coefficient	MAD	MAPE (%)	Pearson's Correlation (Penetration and Average Observed Sharing)
Genre	Last.fm (n=11,150)	29	44	1.5	4	8	0.99
RQ1(a)	Indie & Alternative (n=15,481)	39	51	1.3	4	8	0.99
	Heavy Metal (n=5,319)	32	46	1.4	5	10	0.98
	Trance (n=2,630)	26	43	1.6	6	14	0.98
	Classical (n=3,201)	31	47	1.5	4	18	0.98
	US Survey (n=1,036)	23	32	1.4	5	5	0.98
RQ1 (a)	**Average**	**30**	**44**	**1.5**	**5**	**11**	**0.98**
Artist	Last.fm (n=11,150)	14	29	2	3	12	0.94
RQ1(b)	Indie & Alternative (n=15,481)	25	40	1.6	2	5	0.99
	Heavy Metal (n=5,319)	20	35	1.7	2	5	0.99
	Trance (n=2,630)	17	31	1.8	2	7	0.99
	Classical (n=3,201)	13	28	2.1	2	10	0.98
RQ1 (b)	**Average**	**18**	**33**	**1.8**	**2**	**8**	**0.98**
Album	Last.fm (n=11,150)	7	19	2.7	2	15	0.95
RQ1(c)	Indie & Alternative (n=15,481)	16	31	1.9	2	7	0.99
	Heavy Metal (n=5,319)	11	25	2.2	1	7	0.99
	Trance (n=2,630)	5	12	2.3	2	52*	0.98
	Classical (n=3,201)	5	16	3.2	2	58*	0.97
RQ1 (c)	**Average**	**9**	**21**	**2.5**	**2**	**28**	**0.98**
Song	Last.fm (n=11,150)	7	19	2.8	2	18	0.97
RQ1(d)	Indie & Alternative (n=15,481)	16	34	2.1	2	7	0.98
	Heavy Metal (n=5,319)	11	27	2.4	2	16	0.98
	Trance (n=2,630)	5	15	2.8	2	31*	0.98
	Classical (n=3,201)	5	16	3.5	3	46*	0.94
RQ1 (d)	**Average**	**9**	**22**	**2.7**	**2**	**24**	**0.97**
All	**Average**	**17**	**30**	**2.1**	**3**	**10**	**0.98**

Table 5: Average DoP Tests of Fit for 6-month, 3-month, 1-month and 1-week listening.

Time Period	Penetration Cut-Off (%)	Analysis Level	Average Penetration (%)	Average Observed Duplication (%)	D-Coefficient	MADs	MAPE (%)	Pearson's Correlation (Penetration and Average Observed Sharing)
6-months	≥20% pen	Genre	38	53	1.4	4	8	0.98
RQ2 (a)	≥10% pen	Artist	19	35	1.9	3	11	0.93
	≥10% pen	Album	10	23	2.3	2	10	0.96
	≥10% pen	Song	12	34	2.8	4	14	0.77
RQ2 (a)		**Average**	**19**	**36**	**2.1**	**3**	**11**	**0.91**
3-months	≥20% pen	Genre	35	49	1.4	4	8	0.98
RQ2 (b)	≥10% pen	Artist	16	32	2.0	3	12	0.91
	≥5% pen	Album	9	21	2.6	3	14	0.92
	≥5% pen	Song	8	22	2.8	3	13	0.89
RQ2 (b)		**Average**	**17**	**31**	**2.2**	**3**	**12**	**0.92**
1-month	≥20% pen	Genre	32	46	1.4	3	5	0.98
RQ2 (c)	≥10% pen	Artist	13	31	2.4	4	16	0.79
	≥5% pen	Album	7	19	2.9	4	22	0.58
	≥5% pen	Song	7	30	4.3	5	22	0.88
RQ2 (c)		**Average**	**15**	**31**	**2.7**	**4**	**16**	**0.81**
1-week	≥20% pen	Genre	30	43	1.5	3	6	0.97
RQ2 (d)	≥5% pen	Artist	7	22	3.2	3	25	0.84
RQ2 (d)		**Average**	**18**	**33**	**2.4**	**3**	**16**	**0.90**
All		**Average**	**17**	**33**	**2.4**	**3**	**13**	**0.88**

Table 6: Sharing within and outside of a meta genre for five behavioural datasets (1-year listening)

Data	Classical		Heavy		Indie		Trance		Last.fm		Average	
Genre	Intra	Inter	Intra	Inter	Intra	Inter	Intra	Inter	Intra	Inter	Intra	Inter
Blues			1.6	1.1	1.5	1.1				1.0	1.6	1.1
Classical	0.9	0.9		1.2		1.0		1.1	1.0	1.0	1.0	1.1
Country				1.1		1.0		1.1		1.0		1.1
Electronica/Dance	1.1	1.0	1.6	1.1	1.2	1.0	1.0	1.0	1.0	1.0	1.2	1.1
Folk	1.0	1.0	1.5	1.1	1.0	0.9	1.3	1.1	1.0	1.0	1.2	1.0
Hip-Hop/Rap	1.2	1.0	1.5	1.1	1.0	1.0	1.0	1	1.0	1.0	1.1	1.1
Indie/Alternative	1.0	1.0	1.1	1.1	0.9	1.0	1.0	1.1	1.0	1.0	1.0	1.1
Jazz	1.2	1.1		1.1	1.2	1.1		1.1		1.0	1.2	1.1
Metal	1.5	1.1	1.1	1.0	2.0	1.0	2.3	1.1	1.0	1.0	1.6	1.0
Pop	1.2	1.1	1.3	1.1	1.0	1.0	1.3	1.1	1.0	1.1	1.2	1.1
R&B/Soul/Funk	1.3	1.0	1.7	1.1	1.2	1.0	1.2	1.0	1.0	1.1	1.3	1.1
Rock	1.1	1.1	1.1	1.1	1.0	1.0	1.4	1.2	1.0	1.1	1.1	1.1
Soundtracks	0.9	0.9	0.9	0.9	1.0	1.0	1.1	1.0	1.0	0.9	1.0	1.0
Average	1.1	1.0	1.3	1.1	1.2	1.0	1.3	1.1	1.0	1.0	1.2	1.0
Overall	1.2 Intra PSI						1.0 Inter PSI					

References

Agostini JM. 1961. How to estimate unduplicated audiences. *Journal of Advertising Research,* **1** (3): 11-14.

Agostini JM. 1962. Analysis of magazine accumulative audiences. *Journal of Advertising Research,* **2** (1): 24-27.

Anesbury Z, Nguyen Y, Bogomolova S. 2018. Getting a ‘sweet’ deal: does healthfulness of a sub-brand influence consumer loyalty? *European Journal of Marketing,* **52** (9-10): 1802-1826.

Anesbury ZW, Bennett D, Kennedy R. 2021. How persistent are duplication of purchase partitions? *Journal of Consumer Behaviour,* **Ahead-of-print** (Ahead-of-print).

Anesbury ZW, Jürkenbeck K, Bogomolov T and Bogomolova S. 2020. Analyzing proprietary, private label, and non-brands in fresh produce purchases. *International Journal of Market Research*: 1-23.

Anschuetz N. 2002. Why a brand's most valuable consumer Is the next one it adds. *Journal Of Advertising Research,* **42** (1): 15-21.

Armstrong JS, Collopy F. 1992. Error measures for generalizing about forecasting methods: Empirical comparisons. *International Journal of Forecasting,* **8** (1): 69-80.

Banelis M, Riebe E, Rungie C. 2013. Empirical evidence of repertoire size. *Australasian Marketing Journal,* **21** (1): 59-65.

Barwise P. 1995. Good Empirical Generalizations. *Marketing Science,* **14** (3): G29-G35.

Barwise P, Ehrenberg A. 1988. *Television and its audience.* Sage Publication*:* London.

Bass FM. 1974. The Theory of Stochastic Preference and Brand Switching. *Journal of Marketing Research,* **11** (1): 1-20.

Bennett D. Year. Mapping the Dynamic Chinese Marketplace- The Case of TV's. *In:* ANZMAC, 2004 Victoria University of Wellington, New Zealand.

Bennett D. 2008. Brand Loyalty Dynamics—China's Television brands come of age. *Australasian Marketing Journal,* **16** (2): 39-50.

Bennett D, Ehrenberg A. Year. A Lot Can Be Revealed by a Little Data: Two Purchase Analysis of Fast Food Buying. *In:* Sylvie Chetty & Brett Collins, eds. ANZMAC conference proceedings, 2001 Auckland, New Zealand. Massey University.

Bennett D, Graham C. 2010. Is loyalty driving growth for the brand in front? A two-purchase analysis of car category dynamics in Thailand. *Journal of Strategic Marketing,* **18** (7): 573-585.

Bennett DR, Ehrenberg A. Year. The Way Australians Buy Fast Food. *In:* Proceedings of the European Marketing Academy Conference, May 2002 Braga.

Bennett DR, Ehrenberg A, Goodhardt G. Year. Two purchase analysis of brand loyalty among petrol buyers. *In:* Aron O'Cass, ed. Australian New Zealand Marketing Academy Conference, 28 November - 1 December 2000 Gold Coast, Queensland. Griffith University, 93-96.

Brewis-Levie M, Harris P. 2000. An empirical analysis of buying behaviour in UK high street womenswear retailing using the Dirichlet model. *The International Review of Retail, Distribution and Consumer Research,* **10** (1): 41-57.

Carter J, Silverman F. 2004. An empirical approach to market partitioning: Application to the cigarette market. *Journal of Targeting, Measurement and Analysis for Marketing,* **12** (4): 366-378.

Chatfield C, Goodhardt G. 1975. Results concerning brand choice. *Journal of Marketing Research,* **12** (1): 110-113.

Chowdhury S, Barker A, Trinh G and Lockshin L. 2021. Using the Dirichlet model to predict how fashion brands compete and grow on eBay. *Journal of Consumer Behaviour,* **3** (4): 357-373.

Cohen J, Lockshin L, Sharp B. 2012. A better understanding of the structure of a wine market using the attribute of variety. *International Journal of Business and Globalisation,* **8** (1): 66-80.

Cohen J, Tataru D. Year. The Structure of the French Retail Wine Market: A Duplication of Purchase Approach. *In:* 6th AWBR International Conference, 2011 France. Bordeaux Management School.
Collins M, Beal V, Barwise P 2003. TV Channel Use Amongst Multi-Channel Viewers. *Report 15 for Corporate Sponsors.* Adelaide: Ehrenberg-Bass Institute for Marketing Science.
Colombo R, Ehrenberg A, Sabavala D. 2000. Diversity in analyzing brand-switching tables: The car challenge. *Canadian Journal of Marketing Research,* **19**: 23-36.
Corkindale D, Romaniuk J, Driesener C 2013. How the duplication of viewing law applies to website visiting and some implications. *ANZMAC.* Auckland, New Zealand.
Danaher PJ, Conroy DM, McColl-Kennedy JR. 2008. Who Wants a Relationship Anyway?: Conditions When Consumers Expect a Relationship With Their Service Provider. *Journal of Service Research,* **11** (1): 43-62.
Dawes J. 2008. Regularities in buyer behaviour and brand performance; The case of Australian beer. *The Journal of Brand Management,* **15** (3): 198-208.
Dawes J, Graham C, Trinh G and Sharp B. 2021. The Unbearable Lightness of Buying. *Journal of Marketing Management,* **Forthcoming**.
Dawes J, Nenycz-Thiel M. 2013. Analyzing the intensity of private label competition across retailers. *Journal of Business Research,* **66** (1): 60-66.
Dawes J, Romaniuk J, Mansfield A. 2009. Generalized pattern in competition among tourism destinations. *International Journal of Culture, Tourism and Hospitality Research,* **3** (1): 33-53.
Delsing MJMH, Ter Bogt TFM, Engels RCME and Meeus WHJ. 2008. Adolescents' Music Preferences and Personality Characteristics. *European Journal of Personality: Published for the European Association of Personality Psychology,* **22** (2): 109-130.
Domingues MA, Gouyon F, Jorge AM, Leal JP, Vinagre J, Lemos L and Sordo M. 2013. Combining usage and content in an online recommendation system for music in the long tail. *International Journal of Multimedia Information Retrieval,* **2** (1): 3-13.
Driesener C, Rungie C. 2022. The Dirichlet model in marketing. *Journal of Consumer Behaviour,* **21** (1): 7-18.
Driesener C, Rungie CE, Banelis M. 2017. If the Model Fits, Use It: Methods and Benchmarks for Evaluating NBD-Dirichlet Goodness-of-Fit. *Australasian Marketing Journal,* **25** (4): 288-293.
Dubinsky AJ, Hansen RW. 1982. Improving marketing productivity: the 80/20 principle revisited. *California Management Review,* **25** (1): 96-105.
Dunn PG, de Ruyter B, Bouwhuis DG. 2012. Toward a better understanding of the relation between music preference, listening behavior, and personality. *Psychology of Music,* **40** (4): 411-428.
Ehrenberg A. 1966. Laws in marketing: A tail-piece. *Applied Statistics London,* **15**: 257-267.
Ehrenberg A. 1975. The Structure of an Industrial Market: Aviation Fuel Contracts. *Industrial Marketing Management,* **4**: 273-285.
Ehrenberg A. 1988. *Repeat-buying: Facts, theory and applications.* Oxford University Press*:* London.
Ehrenberg A. 1991. New Brands and the Existing Market. *Journal of the Market Research Society,* **33** (4): 285-299.
Ehrenberg A, Bound J 1999. Customer retention and switching in the car market. *Report 6 for Corporate Sponsors.* Adelaide: Ehrenberg-Bass Institute for Marketing Science.
Ehrenberg A, Bound JA. 2000. Turning data into knowledge, in *Marketing research: State of the art perspectives (The handbook of the American Marketing Association and the Professional Market Research Society),* Chuck Chakrapani (ed.) American Marketing Association*:* Chicago, IL; 23-46.
Ehrenberg A, Goodhardt G. 1969. Practical applications of the duplication of viewing law. *Journal of the Market Research Society,* **11** (1): 6-24.

Ehrenberg A, Goodhardt G. 1970. A model of multi-brand buying. *Journal of Marketing Research,* **7**: 77-84.
Ehrenberg A, Goodhardt G, Barwise TP. 1990. Double Jeopardy revisited. *Journal of Marketing,* **54** (3): 82-91.
Ehrenberg A, Pouilleau B. 1992. The Analysis of a Contingency Table: the Pattern of Car Switching. *Journal of Empirical Generalisations in Marketing Science,* **14**: 1-6.
Ehrenberg A, Uncles MD, Goodhardt GJ. 2004. Understanding brand performance measures: using dirichlet benchmarks. *Journal of Business Research,* **57** (12): 1307-1325.
Fader PS, Schmittlein DC. 1993. Excess behavioral loyalty for high-share brands: deviations from the Dirichlet model for repeat purchasing. *Journal of Marketing Research,* **30** (4): 478-493.
George D, Stickle K, Rachid F and Wopnford A. 2007. The association between types of music enjoyed and cognitive, behavioral, and personality factors of those who listen. *Psychomusicology: A Journal of Research in Music Cognition,* **19** (2): 32-56.
Goodhardt GJ. 1966. Constant in duplicated television viewing. *Nature,* **212** (5070): 1616.
Goodhardt GJ, Ehrenberg ASC. 1969. Duplication of television viewing between and within channels. *Journal of Marketing Research,* **6** (2): 169-178.
Goodhardt GJ, Ehrenberg ASC, Collins MA. 1975. *The television audience: Patterns of viewing.* Gower Publishing: Aldershot, United Kingdom.
Gorard S. 2005. Revisiting a 90-year-old debate: The advantages of the mean deviation. *British Journal of Educational Studies,* **53** (4): 417-430.
Grasby A, Corsi A, Dawes J, Driesener C and Sharp B. 2022. How loyalty extends across product categories. *Journal of Consumer Behaviour,* **21** (1): 153-163.
Hand C, Singh J. 2014. Segmenting the betting market in England. *International Journal of Market Research,* **56** (1): 111-127.
International Federation of the Phonographic Industry 2019. Global Music Report: State of the Industry. London, United Kingdom: IFPI.
Kahn B. 1995. Consumer variety-seeking among goods and services:: An integrative review. *Journal of Retailing and Consumer Services,* **2** (3): 139-148.
Kalwani MU, Morrison DG. 1977. A Parsimonious Description of the Hendry System. *Management Science,* **23** (5): 467-477.
Kelly K. 2008. 1,000 True Fans. *The Technium*.
Keng KA, Ehrenberg A. 1984. Patterns of store choice. *Journal of Marketing Research,* **21** (4): 399-409.
Keng KA, Uncles M, Ehrenberg A and Barnard N. 1998. Competitive brand-choice and store-choice among Japanese consumers. *Journal of Product & Brand Management,* **7** (6): 481 - 494.
Krause AE, North AC, Hewitt LY. 2015. Music-listening in everyday life: Devices and choice. *Psychology of Music,* **43** (2): 155-170.
Lam D. 2006. Applicability of the Duplication of Purchase Law to Gaming. *UNLV Gaming Research & Review Journal,* **10** (2): 55-62.
Lam D, Ozorio B. 2013. Duplication of Purchase Law in the gaming entertainment industry—A transnational investigation. *International Journal of Hospitality Management,* **33**: 203-207.
Last.fm. 2013. Scrobbling 2.0 Documentation. Available at https://www.last.fm/api/scrobbling [Accessed on 24 July 2019].
Last.fm Ltd. 2019. Track My Music. Available at https://www.last.fm/about/trackmymusic [Accessed on 19 November 2019].
Lees G, Wright M. Year. Does the duplication of viewing law apply to radio listening? *In:* Proceedings of the Australian & New Zealand Marketing Academy Conference, 2009 Melbourne, Australia.
Lees G, Wright M. 2013. Does the duplication of viewing law apply to radio listening? *European Journal of Marketing,* **47** (3/4): 674-685.

Lindsay RM, Ehrenberg A. 1993. The design of replicated studies. *The American Statistician,* **47** (3): 217-228.
Mansfield A, Romaniuk J. Year. How do Tourism Destinations Compete? An Application of the Duplication of Purchase Law. *In:* EMAC, 2003 Glasgow. 1-6.
Mansfield A, Romaniuk J, Sharp B 2003. Competition Among International Tourist Destinations: Applying the Duplication of Purchase Law. *ANZMAC.* Adelaide.
McPhee WN. 1963. *Formal Theories of Mass Behavior.* Free Press*:* New York, United States.
Meyer-Waarden L, Benavent C. 2009. Grocery retail loyalty program effects: self-selection or purchase behavior change? *Journal of the Academy of Marketing Science,* **37** (3): 345-358.
Naami T, Anesbury ZW, Stocchi L and Winchester M. 2021. How websites compete in the Middle East: The example of Iran. *Journal of Consumer Behaviour,* **21** (1): 121-136.
Nielsen. 2017. Time with tunes: How technology is driving music consumption. Available at http://www.nielsen.com/us/en/insights/news/2017/time-with-tunes-how-technology-is-driving-music-consumption.html [Accessed on June 2018].
North AC, Hargreaves DJ. 2007. Lifestyle correlates of musical preference: 1. Relationships, living arrangements, beliefs, and crime. *Psychology of music,* **35** (1): 58-87.
Peterson RA, Kern RM. 1996. Changing Highbrow Taste: From Snob To Omnivore. *American Sociological Review,* **61** (5): 900-907.
Rentfrow PJ, Goldberg LR, Levitin DJ. 2011. The Structure of Musical Preferences: A Five-Factor Model. *Journal of Personality and Social Psychology,* **100** (6): 1139.
Rentfrow PJ, Gosling SD. 2003. The do re mi's of everyday life: The structure and personality correlates of music preferences. *Journal of personality and social psychology,* **84** (6): 1236.
Rentfrow PJ, Gosling SD. 2006. Message in a ballad: The role of music preferences in interpersonal perception. *Psychological Science,* **17** (3): 236-242.
Reuver Ad, Jonkheer B 2017. Stop with Why: How a dairy company successfully researched 'how brands grow'. *ESOMAR Conference.* Shanghai.
Ricci F, Rokach L, Shapira B. 2015. Recommender Systems Handbook, in *Recommender Systems Handbook,* Francesco Ricci, Lior Rokach & Bracha Shapira (eds.). Springer*:* New York; 1-36.
Romaniuk J, Dawes J. 2005. Loyalty to price tiers in purchases of bottled wine. *Journal of Product & Brand Management,* **14** (1): 57-64.
Romaniuk J, Sharp B. 2021. *How Brands Grow: part 2.* Oxford University Press*:* Victoria, Australia.
Savage M. 2006. The musical field. *Cultural trends,* **15** (2-3): 159-174.
Scriven J, Bound J, Graham C. 2017. Making sense of common Dirichlet deviations. *Australasian Marketing Journal (AMJ),* **25** (4): 294-308.
Scriven J, Danenberg N 2010. Understanding How Brands Compete: A Guide to Duplication of Purchase Analysis. *Report 53 for Corporate Sponsors.* Adelaide: Ehrenberg-Bass Institute for Marketing Science.
Scriven JA, Ehrenberg A 1994. The Structure of the UK Petrol Market. Reports prepared for Shell International Petroleum Co. and BP Oil Europe.
Sharp B. 2010. How brands grow, in *How Brands Grow,* Byron Sharp (ed.) Oxford University Press*:* South Melbourne, Australia; 16-27.
Sharp B, Romaniuk J. 2016. How brands grow, in *How Brands Grow: Part 2,* Jenni Romaniuk & Byron Sharp (eds.). Oxford University Press*:* South Melbourne; 1-22.
Sharp B, Sharp A 1997. Positioning & Partitioning. *European Marketing Academy Conference.* Coventry, UK: The University of Warwick.
Sharp B, Wright M, Goodhardt G. 2002. Purchase loyalty is polarised into either repertoire or subscription patterns. *Australasian Marketing Journal,* **10** (3): 7-20.
Sjostrom T, Corsi AM, Driesener C and Chrysochou P. 2014. Are food brands that carry light claims different? *Journal of Brand Management,* **21** (4): 325-341.

Sørensen K, Van den Broucke S, Fullam J, Doyle G, Pelikan J, Slonska Z and Brand H. 2012. Health literacy and public health: A systematic review and integration of definitions and models. *BMC public health,* **12** (80): 1-13.
Stern P, Hammond K. 2004. The Relationship Between Customer Loyalty and Purchase Incidence. *Marketing Letters,* **15** (1): 5-19.
Trinh G. 2014. Predicting variation in repertoire size with the NBD model. *Australasian Marketing Journal,* **22** (2): 111-116.
U.S. Census Bureau. 2018. QuickFacts. Available at https://www.census.gov/quickfacts/fact/table/US/AGE295218#AGE295218 [Accessed].
U.S. Census Bureau Population Division. 2017. Estimates of the Total Resident Population and Resident Population Age 18 Years and Older for the United States, States, and Puerto Rico. [Accessed on 24 July 2019].
U.S. Census Bureau Population Division. 2019. Annual Estimates of the Resident Population by Single Year of Age and Sex for the United States: April 1, 2010 to July 1, 2018. Available at https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=PEP_2018_PEPSYASEXN&prodType=table [Accessed on 24 July 2019].
Uncles, Ehrenberg A. 1990a. The Buying of Packaged Goods at US Retail Chains. *Journal of Retailing,* **66** (3): 278-296.
Uncles, Kwok S. 2009. Patterns of store patronage in urban China. *Journal of Business Research,* **62** (1): 68-81.
Uncles MD, Ehrenberg A. 1990b. Industrial Buying Behavior: Aviation Fuel Contracts. *International Journal of Research in Marketing,* **7** (1): 56-68.
Uncles MD, Ellis K. 1989. The Buying of Own Labels. *European Journal of Marketing,* **23** (No. 3): 47-70.
Wilson D, Winchester M. 2019. Extending the double jeopardy and duplication of purchase laws to the wine market. *International Journal of Wine Business Research,* **31** (2).
Winchester M, Lees G. 2013. Do radio stations in New Zealand target successfully? *Australasian Marketing Journal,* **21** (1): 52-58.
Wlömert N, Papies D. 2016. On-demand streaming services and music industry revenues—Insights from Spotify's market entry. *International Journal of Research in Marketing,* **33** (2): 314-327.
Wright M, Sharp A, Sharp B. 1998. Are Australasian brands different? *Journal of Brand and Product Management,* **7** (6): 465-480.
Wright M, Sharp A, Sharp B. 2002. Market statistics for the dirichlet model: Using the juster scale to replace panel data. *International Journal of Research in Marketing,* **19** (1): 81 - 90.
Wrigley N, Dunn R. 1984a. Stochastic panel-data models of urban shopping behaviour: 3. The interaction of store choice and brand choice. *Environment and Planning A,* **16** (9): 1221-1236.
Wrigley N, Dunn R. 1984b. Stochastic Panel-Data Models of Urban Shopping Behaviour: 2. Multistore Purchasing Patterns and the Dirichlet Model. *Environment and Planning A,* **16**: 759-778.
Yapriady B, Uitdenbogerd AL. Year. Combining demographic data with collaborative filtering for automatic music recommendation. *In:* International Conference on Knowledge-Based and Intelligent Information and Engineering Systems 2005 2005 Berlin. Springer, 201-207.