From melody to monopoly? The economics of differentiation in the music industry revisited

Rouven Beiner

July 16, 2024

Abstract

This study explores the non-linear relationship between song success and differentiation using the weekly Spotify Top 200 charts. Without assuming a specific functional form, a non-parametric Local Linear Least Squares (LLLS) model and hypothesis tests with bootstrap methods are employed. The analysis indicates that differentiation is beneficial, aligning with a scenario in a Hotelling model where the market is not in equilibrium. However, the null hypothesis that this effect is zero could only be rejected in one of the two applied non-parametric hypothesis tests, and it is important to note that the effect size is small and likely lacks economic significance. I highlight and discuss the importance of considering temporal paradigms, and the problems of using rank-based measures in the context of highly skewed music industry revenues.

Contents

List of Figures

List of Tables

1 Introduction

As listener habits change, trends evolve, and culture shifts, so do the determinants driving albums to the top of the charts, ranging from melodies and rhythms to lyrics and overall audio characteristics. Understanding these determinants holds significance for several reasons. From a historical and cultural standpoint, the types of music that achieve mainstream success often mirror societal trends and sentiments. Mauch et al. (2015) for example focus on the Billboard top 100 charts from 1960 to

2010, investigating the evolution of popular music in the USA and Terry F. Pettijohn and Donald F. Sacco (2009) examined the relationship between song characteristics of the Billboard Top 100 charts and US economic conditions.

From an economic perspective, profit-oriented record labels and artists can benefit from identifying and adapting to these success determinants. According to estimates by Denisoff (1986), only one in ten albums released by artists under contract is profitable. Publishing music that aligns with characteristics of success can potentially increase its likelihood of reaching a large audience and can help reduce the risk of releasing unprofitable albums or singles. In 1988, the successful producer duo "The KLF" published a book titled "The Manual (How to Have a Number One the Easy Way)." Although this step-by-step guide may not be based on empirical research, it underscores the notion that the success of songs and artists is not purely a matter of chance but can be strategically influenced. Furthermore, it suggests that popular songs may share a kind of identical sound DNA.

The effect of musical attributes of songs are examined by a number of studies. Myra et al. (2018) used UK chart data from 1985 to 2015 to analyze trends and predict success in contemporary songs. Other authors conducted survival analysis (Asai (2008),Bhattacharjee et al. (2007),Giles (2007)) on the number of weeks a song is in the charts. These studies suggest that some specific characteristics of the audio material indeed can have a significant effect on chart success.

However, audio features alone account for only a minor share of the overall variability in chart success. One problem of analysing chart songs is that the set of chart songs is already a sample of the most successful songs, and variation within the charts cannot be described by audio features alone.

Analyzing socioeconomic factors such as promotion budget or social and cultural capital of the musicians (Pinheiro and Dowd (2009), Wapnick et al. (1997)), which are usually hard to measure, can partially fill this gap.

Apart from literature that examined the observable and measurable socioeconomic and musical factors, a broad literature analyzed the music industry in terms of so called superstar dynamics or superstardom, first introduced by Marshall (1947) and Rosen (1983) which exists if "...relatively small numbers of people earn enormous amounts of money, and seem to dominate the fields in which they engage." (Rosen, 1983, p.449). The stardom hypothesis by Marshall (1947) and Rosen (1983) states that small differences in talent or quality explain big differences in success. For the music recording industry this hypothesis was rejected by Hamlen (1991). The other line of explaining superstardom considers a stochastic approach, initially introduced by Chung and Cox (1994). According to the theory, apart from artists related factors, superstardom is a result of luck and arising bandwaggon effects. There is substantial literature on superstardom in the music industry, with extensive research debating which exponential distribution best fits music revenues. To name just a few, Chung and Cox (1994),Strobl and Tucker (2000), Fox and Kochanowski (2004), Giles (2006), Spierdijk and Voorneveld (2009),Fox and Kochanowski (2004) and Cox et al. (1995). However, regardless of the specific distribution type, all power-law distributions (e.g., Pareto, Yule, Lotka) are characterized by long tails and hard-to-predict outliers. Some of these distributions don't even have a finite mean, making OLS predictions effectively useless.

The fourth explanation is that chart success can be explained by differentiation. The basic economic theory of differentiation claims that by creating a heterogeneous good and thus differentiating from competitors, the producer can increase demand by increasing market power and yielding monopolistic profits. Therefore, when analyzing determinants of success in music, apart from the analysis of audio features, socioeconomic factors and the stochastic properties of the income distribution, the audio features of related songs and their similarity to the respective song should be considered. Without analyzing the effect of differentiation, we could find that songs with audio feature XY usually perform better than others. But it is possible that this effect is observed only because the competition in this type of music (characterized by audio feature XY) is low. This could be the case of new emerging trends, where high demand meets low supply. Usually artists who set new trends or implement new trending styles and thus differentiate from others can achieve high chart positions.

In their study, Askin and Mauskapf (2017) proposed that the effect of differentiation can be applied to the music industry and analyzed whether the similarity of a song to its peer group can explain Billboard chart success. They analyzed Billboard top 100 chart data together with songs characteristics as controls. Alexander (1996) also analyzed differentiation in music. In their study, Askin and Mauskapf (2017) found a u-shaped non-linear effect using a ordered logit with a squared term, suggesting that differentiation has an optimum and is beneficial up to a cosine similarity value of 0.54.

I aim to replicate this analysis by examining the weekly Spotify Top 200 song charts using similar variables to those employed by Askin and Mauskapf (2017). I also believe in the non-linearity of the underlying relationship between success and differentiation, but I prefer not to assume a specific functional form. Instead, I apply non-parametric methods to estimate marginal effects (ME) and test hypotheses.

A first argument in favor of non parametric methods such as Local linear least squares (LLLS) is that the underlying data relationship may be more complex than what can be captured by a second-order polynomial function. And secondly, finding a U-shaped relationship in a model with an added squared term is unsurprising and does not inherently provide evidence for a true U-shaped relationship. In a parametric setup, it is hard to prove if the relationship is genuinely U-shaped and if the coefficients are significantly different from zero. Non-parametric methods can capture for the U-shaped relationship without such assumptions, and if the relationship holds true, relying on parametric methods becomes more justified.

The paper is structured as follows: First, I discuss the economic theory of differentiation. Next, I describe the data used, provide descriptive statistics, and conduct an exploratory data analysis. The model-based analysis begins by estimating a bivariate baseline LLLS model, followed by an multivariate model that includes control variables. These models are estimated using similarity metrics for the past 26, 52, and 104 weeks preceding each song. Non-parametric hypothesis testing is then employed to determine whether the effects significantly differ from zero across all values of similarity. Additionally, I analyze the effect of differentiation for specific unique artists.

2 The economic theory of differentiation applied to the music industry

From an economic perspective, the music streaming market is interesting because artists cannot compete on price since the price for a music stream is fixed and cannot be influenced by artists. Therefore the famous bertrand paradox is not present. This implies that even in the case of homogenous goods (i.e., all songs being roughly the same), artists still earn profits.

Without pricing strategies available, the primary way for artists to maximize profits is through differentiation. Artists can differentiate their music both vertically and horizontally. Vertical differentiation involves differentiation with respect to the quality, which can be influenced by major label involvement (usually create professional working environments), collaboration with professional composers or songwriters, and the musical education of the artist. Horizontal differentiation, on the other hand, concerns the characteristics and features of the musical product. As a starting point I apply the symmetric spatial differentiation model employed by Hotelling (1929) to the music industry, illustrating how artists determine the degree of their horizontal differentiation. Figure 1 illustrates basic product differentiation

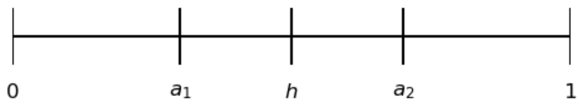

Figure 1: One dimentional hotelling model of differentiation

in a Hotelling model with only two artists songs, where a song's characteristic (such as danceability) ranges from zero to one. A single consumer preference is represented at point h while the two artists' songs are located at points a_1, a_2. The products are homogeneous if $a_1 = a_2$ and differentiated if $a_1 \neq a_2$. A consumer at point h chooses a_1 if:

$$p_1 + t(h - a_1)^2 \leq p_2 + t(a_2 - h)^2 \tag{1}$$

with p_1, p_2 representing the prices, and $t(h - a_1)^2$ and $t(a_2 - h)^2$ representing costs or sacrifices the consumer has to make for deviating from their preference. In the streaming market, the streaming price for consumers is the same for every song $p_1 = p_2$. Consequently, the equation determining whether the consumer chooses a_1 simplifies to:

$$t(h - a_1)^2 \leq t(a_2 - h)^2 \tag{2}$$

When solving equation (2) for h, we find that if the preference is less than or equal to the mean differentiation of both products $\bar{a}$, then the consumer will choose a_1:

$$h \leq \frac{1}{2}(a_1 + a_2) = \bar{a} \tag{3}$$

The result is expected, as it simply means that the consumer chooses the song that best matches the preference (i.e., the closest to the preference point h). In

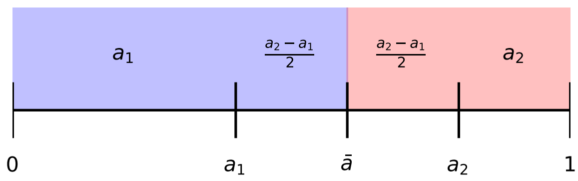

Figure 2: One dimentional hotelling model of differentiation

Hotelling's original model, a large number of consumers are uniformly distributed across the range $a \in (0, 1)$ (Figure 2). This creates an incentive for artists to move toward the center to increase their market share from the edges to their own position, thereby maximizing streams. If both artists are in the equilibrium position at $a = 0.5$ (indicating an optimized market), they have no incentive to deviate, as any deviation would result in a reduction in their streams by $\frac{1}{2}\Delta a$. However, the equilibrium point (0.5,0.5) for both songs (a_1, a_2) is not in the public interest. The welfare optimum would be at the two points (0.25,0.75). Hotelling (1929) states: "Even if A, the first in the field, should settle at one of these points, we have seen that B upon his arrival will not go to the other, but will fix upon a location between A and the centre and as near A as possible. Thus some customers will have to transport their goods a distance of more than $\frac{1}{2}l$, whereas with two stores run in the public interest no shipment should be for a greater distance than $\frac{1}{2}l$." (Hotelling (1929), p. 53). This "Principle of Minimum Differentiation" implied by the Hotelling model is criticized because on the other hand, firms generally aim to maximize product differentiation to reduce price competition as demonstrated by d'Aspremont et al. (1979) and d'Aspremont et al. (1983).

In reality, uniformly distributed consumer preferences in music is not realistic. However, artists can still optimize their positions by moving toward the median of any other distribution of preferences. This concept can also be extended to a multidimensional case, as shown in Figure 3. In Figure 3, we see a two dimensional space with variables $x1$ and $x2$ (e.g. danceability and speechiness). The colors represent individual listeners' preferences, forming a distribution with two centers of mass. The larger points indicate the current audio characteristics of two artists. In a two-dimensional characteristics space, Tabuchi (1994) and Neven and Thisse (1990) identified equilibria where the two firms choose to minimize differentiation in one characteristic while differentiation is maximized in the other. In the case of n dimensions, Irmen and Thisse (1998) show that there are n equilibria in which firms choose maximum differentiation along one (the dominant) characteristic and minimum differentiation along the remaining ones. However, this homogenization of $n - 1$ characteristics holds only when firms compete on prices. In the absence of price competition, Hotelling's principle of minimum differentiation can be extended to the multi-dimensional case. Then, artists tend to move toward the middle between both preference centers, i.e. the median of the bivariate distribution, to maximize streams.

In a market with $n > 2$ artists, assuming uniformly distributed preferences and a variation space of lenght 1, in equilibrium the artists will be uniformly distributed with equal spaces to the next artist. The difference to the next neighbor a_{k+1} is then

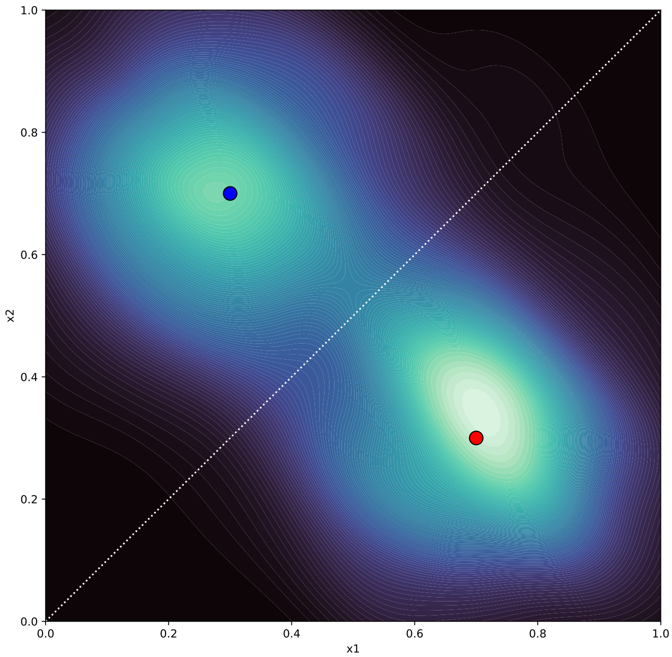

Figure 3: Two dimensional hotelling model of differentiation

$\Delta a = a_{k+1} - a_k = \frac{1}{n}$. In the case of non uniform distributions, the spaces between artists are the negative of the probability density function, meaning at points with higher mass of consumer preferences, the artists are pleaced nearer to each other.

At this point we can conclude that in a world without pricing competition and with perfectly observable listening preferences and profit maxmimizing artists who can change their musical style without costs, there is no incentive to differentiate if the market is already optimized (in equilibrium). Of course at least the assumption of costless repositioning is critical, because artists risk losing their fanbase when changing the musical style. If the market is unoptimized, meaning that artists characteristics are centered around a point other than the optimum, artists have an incentive to differentiate toward the preference mass. This typically occurs when new trends emerge.

Stated differently, the causal relation is as follows: High market optimization means that most songs match the preferences and therefore the optimum. In this scenario, differentiation is expected to have a negative effect because deviating from the other songs also means deviating from the optimum. Thus, when regressing differentiation on rank, the observed effect of differentiation includes two subeffects: the effect of changes in the audio features and the true effect of differentiation.

Consider the following theoretical example with the simple model of the form: $rank = \alpha + \beta * differentiation + \epsilon$. Assume that another variable, *tempo*, has a nonlinear, squared effect on *rank*, with an optimum around 120 BPM. Further assume an optimized market where artists are normally distributed around this 120

BPM optimum. If an artist near the optimum decides to create much faster music at 140 BPM, a negative estimated effect of β in the simple model is expected. In this case, differentiation does not pay off. However, the loss in streams for this artist is not necessarily due to differentiation per se, but rather due to the change in tempo, resulting in music that no longer aligns with consumer preferences. When *tempo* is included as a control variable in the model, the negative effect is captured by the *tempo* variable, and the effect of *differentiation* is no longer biased. Therefore, it is essential to include control variables in the model.

Now, let's consider a market that is not optimized, where most artists are centered around a point that is not the optimum (e.g., 100 BPM). Suppose one artist slows down their music to around 80 BPM, while another changes it to 120 BPM (the optimum). Both artists have the same *differentiation* value, but the second artist can increase streams while the first can not. In this scenario, the coefficient β will likely not be significant. Once again, including control variables will capture this effect, allowing for an accurate interpretation of the *differentiation* effect.

The described effects are essentially endogeneity issues. Omitting variables Z that are correlated with a regressor X and also are correlated with the dependent variable Y causes a correlation between the regressor and the error term $corr(X, \epsilon) \neq 0$, resulting in biased estimates.

3 Data

The Dataset contains the US top 200 weekly Spotify charting Songs during august 2017 to the may 2024 with a total of 7634 unique songs. Audio features and characteristics are derived from the spotify API. The Audio features energy, speechiness, acousticness, instrumentalness, liveness, valence and danceability are computed by spotify and are measures in the range between 0 to 1. The descriptions of these variables in Table 2 are derived from the spotify webpage. The underlying calculation algorithm has not been published; however I think that the 12 Timbre Vectors defined by spotify play a major influence.

The 12 Timbre Vectors are abstract parameters that describe various aspects of a song's timbre, typically centered around zero. Spotify defines these vectors through 12 distinct functions and patterns in the audio spectrogram, each representing a specific tonal characteristic. For example, timbre 1 captures overall volume, timbre 2 represents brightness, and timbre 4 measures attack.

Apart from the audio features, the musical variables tempo (in BPM), duration (in milliseconds) and loudness (in decibel) are in the dataset. Furthermore, I decided not to include artist fixed effects because the characteristics of an artist's song catalog are usually not spread across the whole range of audio features but rather have a specific character. When including artist fixed effects, the effects are 'within group', but I am interested in general effects. To capture some of the artist related effects, especially the superstar effect, I estimated the number of chart entries in the last 52 weeks of the tracks artist.

The variable similarity (as the opposite of differentiation) is the estimated cosine similarity. The Cosine similarity is a measure used to determine how similar two vectors are. It is defined as the dot product of the vectors divided by the product of their magnitudes, or lenghts:

$$similarity(\mathbf{X}, \mathbf{Y}) = \cos(\theta) = \frac{\mathbf{X} \cdot \mathbf{Y}}{\|\mathbf{X}\| \|\mathbf{Y}\|}$$

Where the vector **X** contains the features of the respective song and the vector **Y** containts the features of all charting songs, released in the previous 26 (52,104) weeks before the song. From these previous charting songs all songs from the same artist are excluded. The resulting vector contains similarity values for the similarity of the respective song with each other song previously released and to derive the final similarity value for each song, the mean of these single similarities are estimated. Similarity is estimated with different time horizons (26, 52 and 104 weeks) for the previous songs because I want to check if the effect differ on this dimension.

One challenge lies in determining how success should be measured. Existing literature varies, with some studies employing the highest chart position while others consider the number of weeks an album stays on the charts. But these estimations fail to consider the skewed income distributions in the music industry. For instance, using the highest chart position as a measure implicitly assumes a linear income distribution, as a change from position 200 to 199 is weighted equally to a shift from 2 to 1. Similarly, if the duration on the chart is used, the model cannot differentiate between two albums that remain on the charts for the same duration but occupy vastly different positions. An effective measure of chart performance should account for both the chart position and the duration an album remains on the charts. A simplistic method might involve summing up all chart positions, but this again overlooks the income distribution. In this analysis, I decided to use the peak chart position as an approximation of a song's streaming success.

The spotify chart ranking relies completely on streaming success. Spotify and most major streaming services, use the pro-rata royalty-allocation system, where artists receive a fixed amount for each stream of at least 30 seconds from a user. The value of a stream is calculated by dividing the total subscription revenue by the number of streams in a given period, making streams a direct proxy for streaming revenue.

But is it also appropriate to approximate general monetary success with streaming revenue? It is important to note that an artist's income depends on several sources, including income from physical and digital sales, live performances, merchandise sales and music streaming. But streaming revenues contribute a significant share of total income, accounting for 67% of industry revenues worldwide and even more in Europe and the US (International Federation of the Phonographic Industry (2024)). Since I only observe streaming revenues, there may be a bias in the data because the share of streaming income from all income sources of a musician may vary significantly across different genres. Consequently, when only observing streams, songs from genres where streaming is less important, will perform worse in streaming charts compared to a chart ranking that considers all income streams. Therefore, while this analysis can clearly identify streaming success, and the effect of differentiation on streaming success, generalizing it to overall success may be biased.

Table 1: Detailed description of variables

Variable	**Description**
Artist popularity	The artists popularity measured by the number of charting albums in the previous 52 weeks.
Danceability	Danceability describes how suitable a track is for dancing.
Energy	Energetic tracks feel fast, loud, and noisy (e.g. death metal).
Loudness	The overall loudness of a track in decibels (dB)
Speechiness	Speechiness detects the presence of spoken words in a track. (e.g. talk show, audio book, poetry).
Acousticness	A confidence measure from 0.0 to 1.0 of whether the track is acoustic.
Instrumentalness	Detects whether a track contains no vocals.
Liveness	Detects the presence of an audience in the recording.
Valence	The tracks valence. Tracks with high valence sound more positive (e.g. happy, cheerful, euphoric), tracks with low valence sound more negative (e.g. sad, depressed, angry).
Tempo	The overall estimated tempo of a track in beats per minute (BPM).
Duration	The tracks duration in minutes.
Similarity 26 / 52 / 104	The cosine similarity of the song compared to other charting songs during the previous 26 / 52 / 104 weeks.

4 Descriptive Statistics

Figure 4 displays boxplots of the rank variable grouped by similarity values. The wider boxplots in the similarity ranges from 0.8 to 1 indicate that most songs similarity values are in this range. On first sight the median values as well as the quantiles reveal no clear positive or negative relationship between similarity and rank.

Figure 5 plots the similarity values of hits (ranks 1-10) and non hits (ranke 11-200) for each year separately. Interestingly, three distinct paradigms emerge: (1) In 2017 and 2018, hit songs exhibit significantly higher similarity values than non-hits. (2) From 2019 to 2021, the trend reverses, with hit songs displaying lower similarity values on average (3) From 2022 onward, hit songs once again show higher similarity values. These observations suggest the necessity of splitting the dataset into these three phases for separate analysis. Additionally, a downward trend in similarity is evident from 2017 to 2022, followed by an increase in similarity for both hits and non-hits. This downward trend corresponds with the findings of Askin and Mauskapf (2017). However, the ranges of different similarity values are very amll, ranging from 0.875 to 0.905.

Figure 6 displays the distributions of the three similarity estimations, with hits shown on the left and non-hits on the right. All three distributions exhibit similar properties and moments. When comparing hits and non-hits, it is apparent that the distributions for hit songs have more probability mass in the tails. Overall, the tails are very long, indicating the presence of extreme outliers.

Figure 7 shows the violin plots for the seven audio features. On first sight

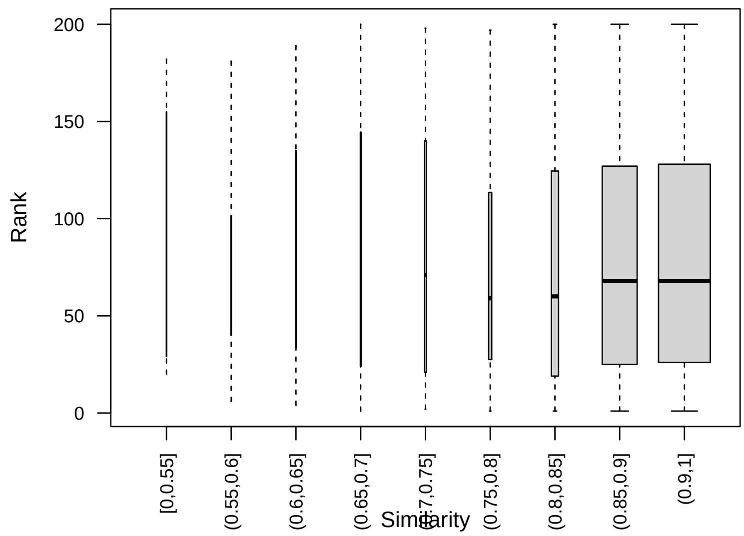

Figure 4: Boxplots grouped by similarity

the distributional properties for hits are the same as for non hits. The variables *instrumentalness*, *energy* and *valence* exhibit a symmetric distribution while the variables *instrumentalness*, *liveness*, *acousticness* and *speechiness* are skewed to the right. Figure 8 shows the distributional properties of the variables *loudness*, *duration*, *artistpopularity* and *tempo*. The left axis scales is for the variable *loudness*, *duration*, *artistpopularity* and the right axis scale is for *tempo*.

To assess multicollinearity issues, I estimated a non-parametric VIF value for each regressor. The non-parametric VIF is derived from the R^2 of a LLLS model for each regressor. For the VIF estimation, each variable was scaled to the $(0, 1)$ range, and a default bandwidth of 0.15 was used. Table 2 presents the VIF values for each regressor. None of the VIFs exceed 5, indicating that multicollinearity is likely not a concern. It is also important to note that non-parametric VIF values tend to be higher than standard OLS VIF values, as the R^2 in an LLLS model is typically higher compared to the OLS estimator.

5 Statistical model and tests

I assume that the effects of variables and differentiation on chart success are non-linear for at least some variables. For instance, tempo probably does not have a linear effect, as a positive linear effect would imply that artists should produce music with infinitely high tempos, while a negative linear effect would suggest a tempo of zero. Due to the complex effects, a local regression with smaller weights for points farther from x is appropriate. Local regressions minimize the sum of weighted least squares, this includes standard OLS as the special case where all the weights are 1. The Local Linear Least Squares (LLLS) method is superior to the

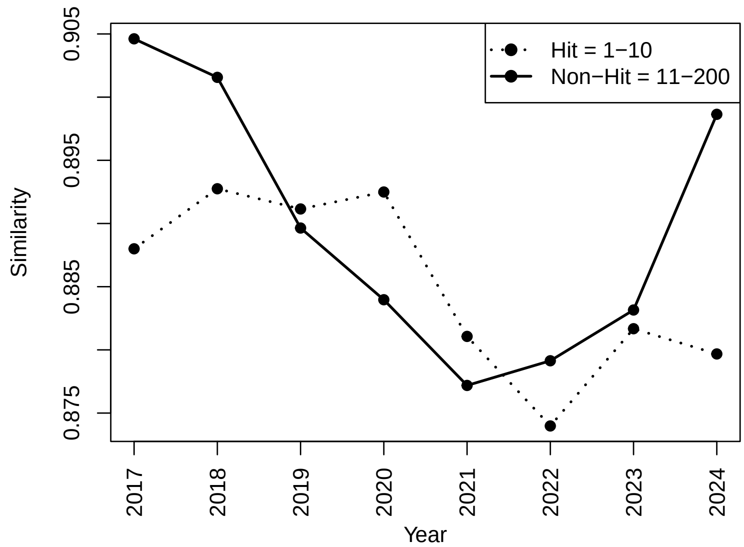

Figure 5: Similarity by year and chart rank

Local Constant Least Squares (LCLS) regression because it has no bias if the true model is linear. Unlike LCLS, LLLS is unbiased at the endpoints.

First, I estimate a baseline model using only similarity as the regressor. To address potential endogeneity, I include audio features as control variables in a second, multivariate model. Including these control variables is further validated by comparing the R^2 of the bivariate model (0.00007) and the multivariate model (0.624). The issue of outliers should be addressed as they appear to significantly influence the similarity analysis. Therefore, a model incorporating outlier removal should be considered. Both models, the bivariate and the univariate are estimated on the whole dataset and the dataset with outlier removal. Outliers in similarity are removed using the interquartile range method: $IQR = Q3 - Q1, \quad \text{Lower Bound} = Q1 - 1.5 \times IQR, \quad \text{Upper Bound} = Q3 + 1.5 \times IQR$. The estimated lower bound is 0.882 and the upper bound is 0.972.

The LLLS estimation for a specific vector x is defined as:

$$\min_{\beta_1(\mathbf{x}),\ldots,\beta_P(\mathbf{x})} \sum_i \left(Y_i - \sum_{p=1}^{P} \beta_p(\mathbf{x}) X_{p,i} \right)^2 k(\mathbf{X}_i, \mathbf{x}, \mathbf{H}), \tag{4}$$

The minimization problem estimates the optimal parameters β_p for each variable $p = 1...P$, such that the distance of Y_i to $\beta_p \mathbf{X}_{p,1}$ for each variable and each datapoint $i = 1...N$ is minimized. The model includes an intercept $X_{1,i} = 1$ for all i, caputured by β_1. The kernel function is defined by:

$$k_{pN}(\mathbf{X_i}, \mathbf{x}, \mathbf{h}) = K_N \left(\frac{|X_{1,i} - x_1|}{h_1} \right) \times K_N \left(\frac{|X_{2,i} - x_2|}{h_2} \right) \times \cdots \times K_N \left(\frac{|X_{P,i} - x_P|}{h_P} \right), \tag{5}$$

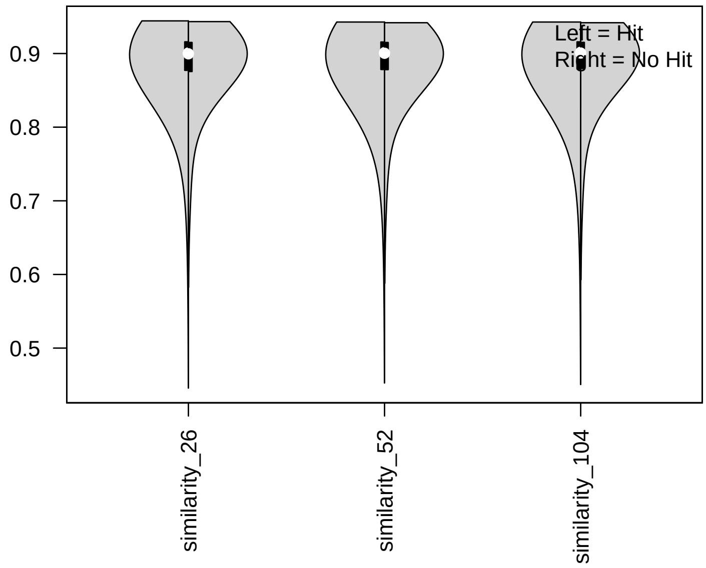

Figure 6: Boxplots of similarity estimates by chart rank

and takes as parameters the bandwidth H and the distribution type. The distribution type to specify how data points are weighted is a gaussian product kernel:

$$K_N\left(\frac{|X_{1,i}-x_1|}{h_1}\right)=\frac{\exp\left\{-\frac{1}{2}\left(\frac{|X_{1,i}-x_1|}{h_1}\right)^2\right\}}{\sqrt{2\pi}} \tag{6}$$

To find the parameter vector $\hat{\beta_p(\mathbf{x})}$, take the partial derivative of (4) with respect to $\beta_q(x)$ for $q=1,\ldots,P$ and set it to zero:

$$\frac{\partial}{\partial\beta_p(x)}\sum_i\left(Y_i-\sum_{p=1}^{P}\beta_p(x)X_{p,i}\right)^2 k(X_i,x,H)=0, \tag{7}$$

$$-2\sum_i\left(Y_i-\sum_{p=1}^{P}\beta_p(x)X_{p,i}\right)X_{p,i}k(X_i,x,H)=0, \tag{8}$$

Solving for $\beta_p(x)$:

$$\sum_i Y_iX_{p,i}k(X_i,x,H)=\sum_i\sum_{p=1}^{P}\beta_p(x)X_{p,i}^2k(X_i,x,H), \tag{9}$$

$$\hat{\beta}_p(x)=\frac{\sum_i Y_iX_{p,i}k(X_i,x,H)}{\sum_i X_{p,i}^2k(X_i,x,H)}. \tag{10}$$

The solution to the minimization problem is then given by:

$$\hat{\beta}(x)=(X'W_xX)^{-1}X'W_xY, \tag{11}$$

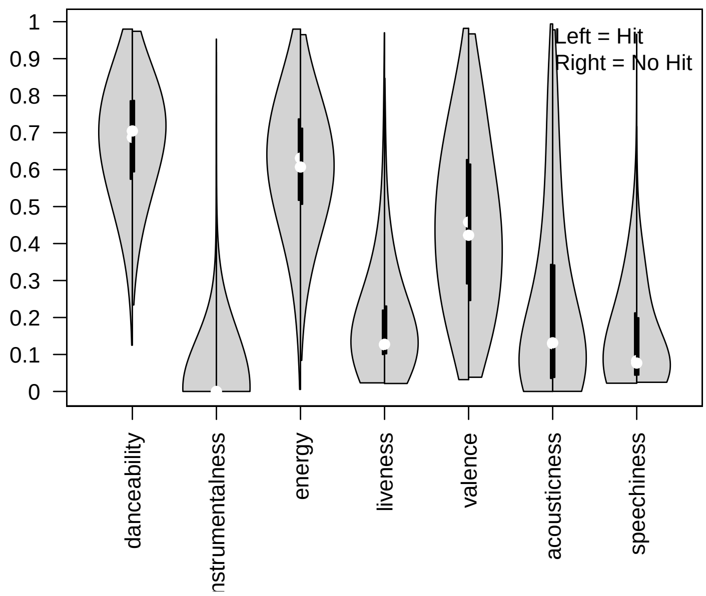

Figure 7: Boxplots of audio features by chart rank

where

$$W_x = \begin{pmatrix} K\left(\frac{X_1-x}{h}\right) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & K\left(\frac{X_n-x}{h}\right) \end{pmatrix}.$$

The same marginal effects can be achieved by estimating the differences of the predicted values $\hat{Y}$ with small h:

$$\hat{\beta}_p(x) = \frac{\hat{Y}(x_1, \ldots, x_p + h_p/2, \ldots, x_P) - \hat{Y}(x_1, \ldots, x_p - h_p/2, \ldots, x_P)}{h_p}, \qquad (12)$$

Choosing the Bandwidth in a LLLS model has a great impact because a small bandwidth will reduce bias but introduce overfitting and a high variance. On the other hand a high bandwidth will create a low variance estimator, in extreme case a linear one, with possible failing to capture the true non linear relationship. Optimal bandwidths are estimated by regressing each predictor in a simple local linear least squares (LLLS) model on the target variable *rank* for each bandwidth using 10-fold cross-validation. The bandwidth optimization is performed on both the whole dataset and the dataset with outlier removal. Detailed bandwidth estimation procedures and visualisations are provided in the appendix.

Investigating the effect of the similarity variable is of primary interest Predictions, as well as marginal effects, are estimated at 100 design points in the range $(similarity_{min}, similarity_{max})$, holding all other variables constant at their mean value.

To estimate the variability of the predicted values, confidence intervals are constructed using a standard bootstrap method. In this approach, n observations are drawn from the original data with replacement to create bootstrap samples

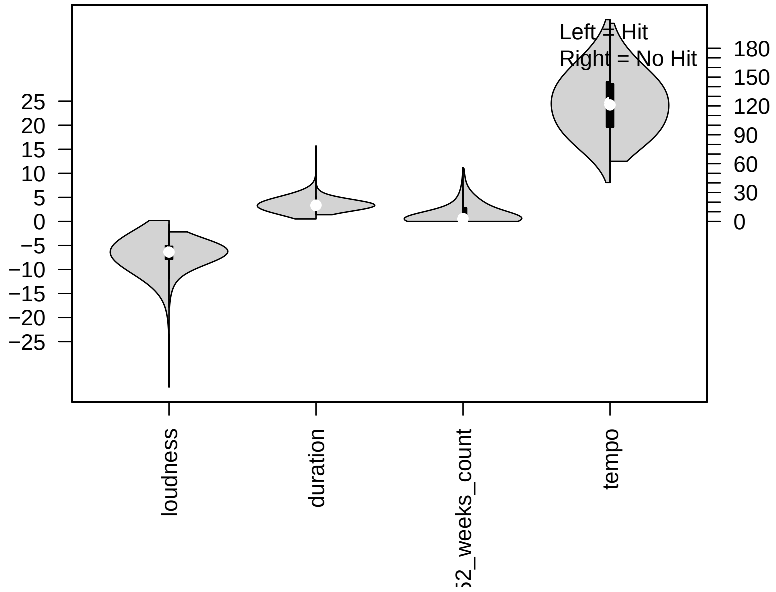

Figure 8: Boxplots of audio features by chart rank

$b = 1, \ldots, B$. For each bootstrap sample, a bootstrap model $LLLS_b$ is estimated to predict values for each design point. These bootstrap predictions are then used to construct the confidence intervals.

The marginal effects are estimated according to equation 12, where the prediction for x_i is subtracted from the prediction for x_{i+1} to obtain the marginal effects. Regarding statistical tests, I conduct two standard tests to determine whether the estimated effect at each design point is significantly different from zero.

Both tests are conducted using resampling methods due to the limitations of conventional methods. The standard testing approach assumes the ability to draw infinitely many samples from a population, compute the statistic of interest (e.g., the t-statistic) for each sample, and then calculate the distribution of the statistic across all samples. However, researchers typically do not have access to the entire population and must make assumptions about the distribution of the test statistic, often assuming a t-distribution.

Consider a scenario where we have estimated a test statistic, such as the t-statistic $\frac{\beta}{SE}$ from our sample and aim to test whether we can reject the null hypothesis that the coefficient is zero. If the variable indeed has no effect, the distribution of the t-statistics from different samples of the population would follow a t-distribution centered around zero. By estimating the percentile of the t-distribution that corresponds to our t-statistic, we obtain the p-value. This p-value describes the probability of observing a test statistic as extreme as the one obtained, assuming the variable has no effect on the target. A small p-value allows us to reject the null hypothesis, indicating that there is an effect.

This procedure is problematic for some reasons. First, in cases where the entire population is included in the dataset, the logic of drawing samples is violated since we have already observed the whole population. This is the case in chart analysis.

Regressor	VIF
danceability	2.26
instrumentalness	2.21
energy	4.16
liveness	1.37
valence	1.69
acousticness	2.39
speechiness	1.78
artist popularity	1.27
loudness	4.1
duration	1.8
tempo	1.37

Table 2: Variance inflation factors (VIF) for regressors

Second, the described procedure relies on strong assumptions regarding the null distribution, and we often do not know the true null distribution. The standard approach to hypothesis testing estimates the standard error (SE) by first calculating the sample standard deviation (SD) and then dividing it by the square root of the sample size. This relies on the central limit theorem to approximate the SE. But if the sample is not representative of the population or contains outliers, the SE estimate can be biased. Third, in large samples, the standard hypothesis testing approach can lead to statistically significant results for even very small effect sizes. This occurs because as the sample size N increases, the standard error ($SE = \frac{SD}{N^{0.5}}$) decreases, causing the t-statistic ($\frac{\beta}{SE}$) to increase. As a result, the p-value decreases, making it easier to reject the null hypothesis. Using resampling procedures can address these issues in most cases.

The null hypothesis is that the marginal effect of similarity on streaming success is zero for every design point. I aim to test this hypothesis and reject it if my estimated test statistic proves to be extremely unusual under the null distribution. For this test, I use two alternative procedures: inverting the confidence interval and imposing the null hypothesis. Both procedures rely on approximating the sample variability with the bootstrap resampling technique. The confidence interval around the gradients can be created by estimating the marginal effects for n bootstraps and then determining the a and $1 - a$ percentiles. The null hypothesis can be rejected if the confidence interval excludes zero.

The second approach involves inverting the null distribution. The similarity indices are reshuffled while all other variables are hold in their original order. Using a bootstrap procedure with $B = 250$ bootstraps, the frequency with which random data would produce a value as extreme as the estimated test statistic is determined.

$$\text{p-value}_i = \frac{\sum_{b=1}^{B} \mathbb{I}\left(|\text{ME}_{b,i}| \geq |\text{ME}_i|\right) + 1}{B + 1}, \quad \text{for } i = 1, \ldots, 100$$

with $i = 1, ..., 100$ are the similarity design points, $\mathbb{I}$ is the indicator function which is 1 if the condition inside is true and 0 otherwise, and ME_b is the Marginal effect from the model with the reshuffled similarity indices.

6 Model estimations

6.1 Industry wide general effects

From observing the distribution in Figure 6, it is evident that the distributions for each similarity measure are nearly identical. Further analysis focuses on the similarity estimate with a 52-week time horizon, as used by Askin and Mauskapf (2017). Figure 9 presents the predictions of the first baseline bivariate model

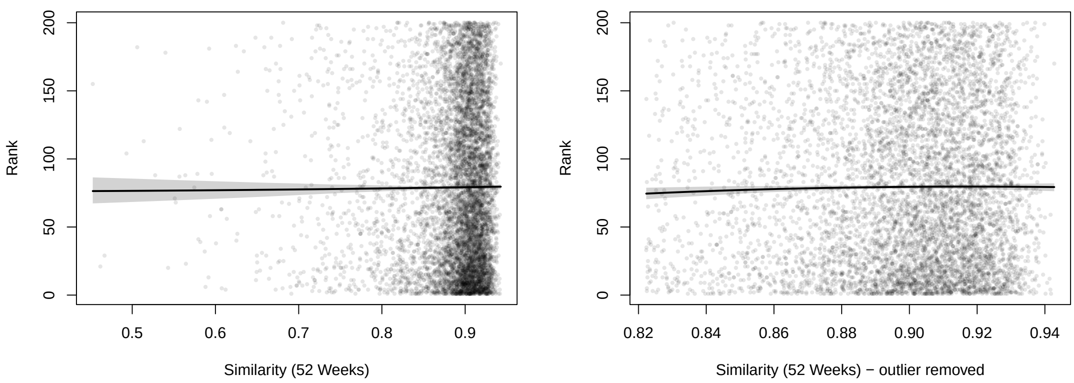

Figure 9: LLLS predictions for *similarity 52* - bivariate model - whole dataset and dataset with outlier removal

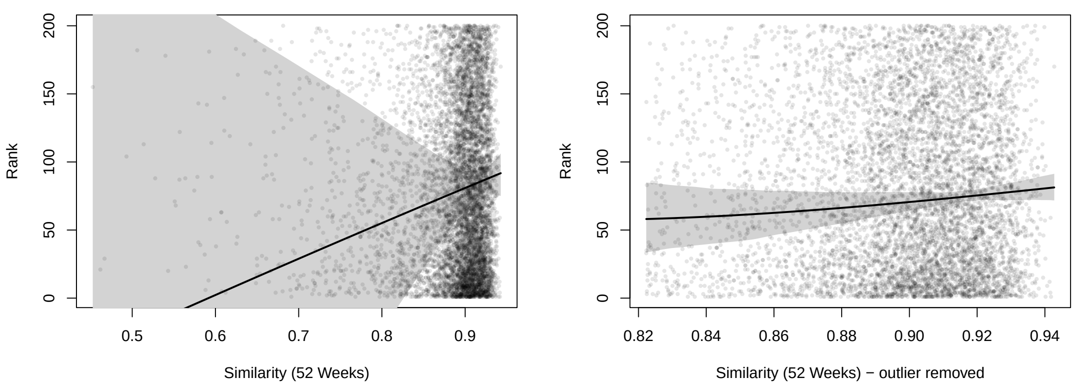

Figure 10: LLLS predictions for *similarity 52* - multivariate model - whole dataset and dataset with outlier removal

along with non-parametric 90% confidence intervals (CIs), both with and without outliers. In Figure 10, the predictions for the multivariate models are shown.

6.1.1 Bivariate models

In both bivariate models, the slope of the predictions is nearly zero, indicating the absence of any significant effect. According to Hotelling's theory, this suggests that

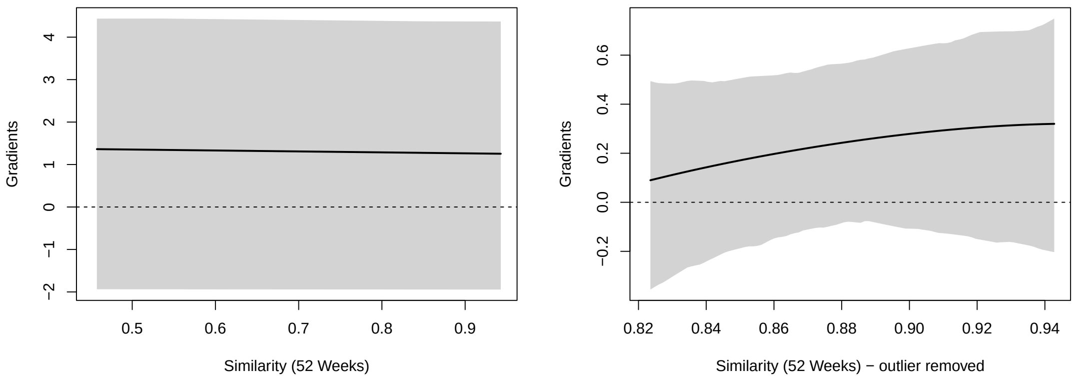

Figure 11: LLLS gradients for *similarity 52* - multivariate model - whole dataset and dataset with outlier removal

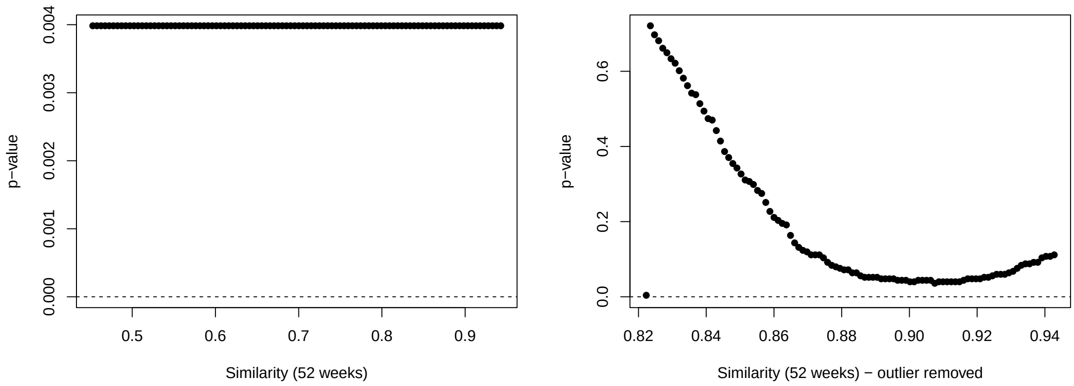

Figure 12: Non parametric p-value for *similarity 52* - whole dataset and dataset with outlier removal

differentiation does not confer an advantage and that the market is in equilibrium.

6.1.2 Full multivariate model

In contrast, the full multivariate model displays a surprisingly strong positive slope (on the left side of Figure 10). However, the wide CIs at the lower end of similarity indicate low accuracy in the predictions, with the upper CI boundary even suggesting a negative slope. The insignificance of the effect can be partly explained by examining the distribution of similarity values. In the lower range of similarity (e.g., below 0.8), there are very few observations, with most data points concentrated around a similarity value of 0.9. This distribution makes it difficult to accurately capture the effect of similarity, leading to insignificance. Nevertheless, this explanation is insufficient, as the CIs of the bivariate model, though increasing, are not as large as those in the multivariate model. The large CIs and marginal effects indicate high variability in predictions for low similarity values, further explained by the estimation method of similarity. Songs with low similarity are often outliers in other variables, resulting in more extreme predictions. This can be seen in Figure 15 in the appendix.

To address this issue, it might be beneficial to consider alternative similarity measures beyond cosine similarity or to remove outliers in similarity, as done in the alternative model. Figure 11 displays the non-parametric marginal effects for the similarity design points along with the corresponding 90% CIs. For the full model, the zero is included within the CIs for each similarity design point, and the effects for each similarity value are not significantly different from zero. Additionally, non-parametric p-values for the probability of observing a value as extreme as the estimated marginal effects (MEs) under the null hypothesis that the effect is zero are estimated. For the full multivariate model, the p-values (Figure 12) indicate that the estimated MEs are unlikely to be achieved with random data, as no ME from the reshuffled data was as extreme as the estimated original ME. However, the relationship between outliers and extreme predictions can explain the large marginal effects that in turn lead to the small p-values.

6.1.3 Multivariate model with outlier removal

The multivariate model with outlier removal exhibits a slightly positive slope, and the CIs are much narrower than in the full multivariate model. The CI of the marginal effects also includes zero for each design point, as seen in Figure 11. The p-values in Figure 12 indicate that MEs at lower similarity values are common in random data, but for similarity values above 0.87, the estimated MEs are unusual compared to random data.

The full model's extreme outliers may bias the estimations, making the model with removed outliers more reliable. The bivariate models provide a good starting point and are important for comparison. However, as explained in Section 3, omitted control variables can lead to endogeneity. Thus, the multivariate model with outlier removal appears to be the most reliable. it is important to note that the threshold for outlier identification is set at a similarity value of 0.882. Consequently, songs with a similarity value below this threshold are excluded from the analysis. This limitation should be taken into account when interpreting the model's results after outlier removal

The logical implication of the slightly positive ME in the outlier removed model is that artists with high similarity values, who are at the center of mass, can increase streams by differentiating from their peer group. According to the Hotelling model, this result suggests that the market is not in equilibrium, possibly due to recent trends that most artists have not yet adopted.

While the non-parametric p-value estimation confirmed that the ME is unlikely to occur if the true effect is zero, the sparse data in the low similarity ranges result in very large CIs, even including zero. This suggests high variability and insignificance. Although one out of two significance tests indicates a significant effect, it is important to note that the effect size is small and likely lacks economic significance.

6.2 Artists' unique effects

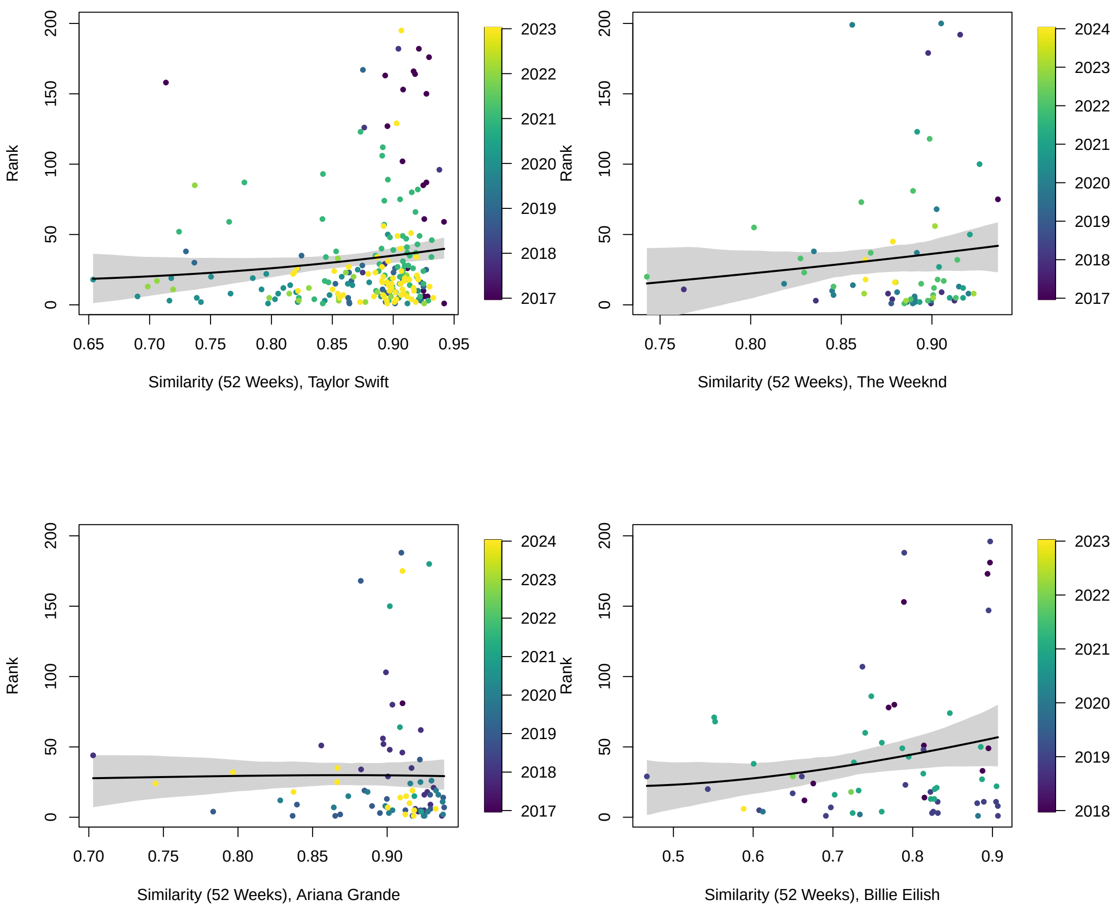

Figure 13: Artist unique predictions for *similarity 52* - bivariate model with bw. 0.2

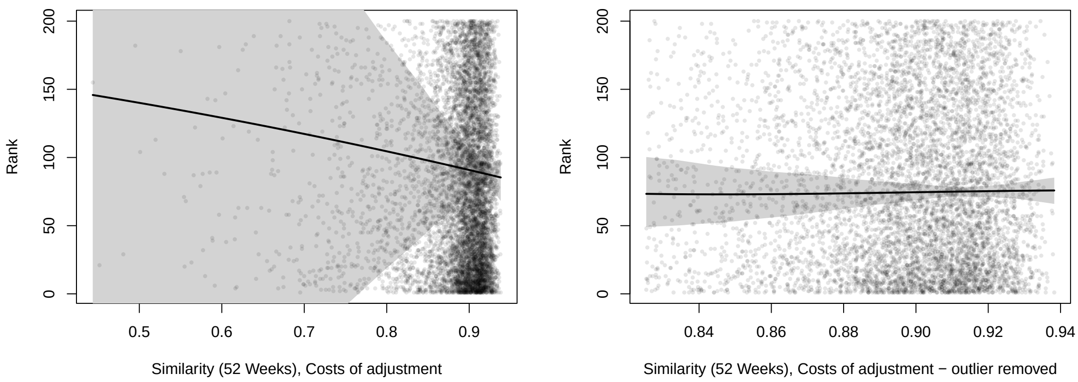

Figure 14: LLLS predictions for *similarity 52* - Adjustment costs - whole dataset and dataset with outlier removal

Until now, the industry-wide general effect of differentiation has been estimated. However, analyzing the effect for individual artists based on their unique song catalogues could provide valuable insights. To estimate artist-specific effects, a bivariate LLLS model without outlier removal is used. Figure 13 presents the predictions for four popular artists—Taylor Swift, The Weeknd, Ariana Grande, and Billie Eilish—along with the year each song reached its highest chart position. Notably, these artists also have songs in the lower similarity range. For all four artists, the songs that perform the worst generally exhibit high similarity. This observation is further confirmed by the positive marginal effects observed for Taylor Swift, The Weeknd, and Billie Eilish. Visual inspection of these four artists' careers does not show that their early songs have lower similarity than those produced in later stages.

As explained in Section 2, adopting new musical trends and changing musical styles can be costly for artists, risking the loss of a loyal fanbase. On the other hand, differentiation is crucial for increasing streams. The costs of adjusting to new musical styles can be estimated empirically by analyzing the similarity values of an artist's songs based on their previously released songs and regressing the similarity values on the chart ranks. The results, presented in Figure 14, indicate that predictions for the outlier removed model are constant, suggesting no effect. This implies that the slightly positive effect of differentiation form Figure 10 and the negative effect of losing a fanbase when adopting new styles cancel each other out.

7 Discussion and conclusion

In cultural industries, the success of products often depends on unpredictable social dynamics, as explained by the superstardom theory. However, numerous studies have demonstrated that product features, such as musical characteristics of the song or socioeconomic attributes of the artist, play a significant role. Moreover, analyzing songs independently is insufficient, the success of a song is interacted with the characteristics of other peer-songs. This is primarily because artists compete for a limited resource: the listening time of music consumers. This paper explores how the economic theory of product differentiation applies to the music streaming market, where price competition is absent. Through a quantitative, non-parametric analysis, I investigated the effect of differentiation on streaming success. The results suggest that differentiation is beneficial, consistent with the situation described in Hotelling's theory where the market is not in equilibrium. However this result does only hold when removing outliers, and the hypothesis that the effect is zero, could only be rejected in one of the two non parametric tests. Additionally, I identified three distinct paradigms within the analyzed period from 2017 to 2024.

A more refined analysis could adress several shortcomings of this study. Consumers preferences and trends are constantly evolving, and analyzing each time period separately, could yield more insightful results. Additionally, each artist's situation is unique. For example, one artist with a loyal fanbase might find that adopting new styles harms their success, while another focus on profit maximization and might adopting new styles when they emerge. Therefore, it would be beneficial to analyze individual artists in greater detail. Regarding the similarity measure, using more variables could better capture similarity and differentiation. Another improvement for further analysis is to estimate similarity for single characteristics. Regarding model specification, a non-parametric F-test for the relevance of each regressor could be conducted. Additionally, the production function specification should not assume perfectly separable and substitutable inputs, as this assumption is unrealistic and is present in the current specification. To account for complementarities between input factors, interaction terms should be included. Using ranks as the endogenous variable in chart analysis is problematic. This approach fails to account for the skewed distribution of streams because the interpretation of the differences between higher ranks (e.g., 199-200) is the same as between lower ranks (e.g., 1-2). Given that income in the music industry is highly skewed, accounting for this in the analysis would provide more accurate insights. Exploring the degree of optimization in the music market could yield valuable insights. This can be achieved by comparing model predictions for a specific variable with the density of released songs for that variable. These predictions would serve as an approximation of consumer preference distribution. In an optimized market, the distribution of songs should align with consumer preference distribution.

References

Alexander, P. J. (1996). Entropy and popular culture: Product diversity in the popular music recording industry. *American Sociological Review*, 61(1):171–174.

Asai, S. (2008). Factors affecting hits in japanese popular music. *Journal of Media Economics*, 21(2):97–113.

Askin, N. and Mauskapf, M. (2017). What makes popular culture popular? product features and optimal differentiation in music. *American Sociological Review*, 82(5):910–944.

Bhattacharjee, S., Gopal, R. D., Lertwachara, K., Marsden, J. R., and Telang, R. (2007). The effect of digital sharing technologies on music markets: A survival analysis of albums on ranking charts. *Management Science*, 53(9):1359–1374.

Chung, K. H. and Cox, R. A. K. (1994). A stochastic model of superstardom: An application of the yule distribution. *The Review of Economics and Statistics*, 76(4):771–775.

Cox, R. A. K., Felton, J. M., and Chung, K. H. (1995). The concentration of commercial success in popular music: An analysis of the distribution of gold records. *Journal of Cultural Economics*, 19:333–340.

d'Aspremont, C., Gabszewicz, J. J., and Thisse, J.-F. (1979). On hotelling's "stability in competition". *Econometrica*, 47:1145–1150.

d'Aspremont, C., Gabszewicz, J. J., and Thisse, J.-F. (1983). Product differences and prices. *Economics Letters*, 11:19–23.

Denisoff, R. S. (1986). *Tarnished Gold: Record Industry Revisited*. Routledge, New York, 1st edition.

Fox, M. A. and Kochanowski, P. (2004). Models of superstardom: An application of the lotka and yule distributions. *Popular Music and Society*, 27(4):507–522.

Giles, D. E. (2006). Superstardom in the us popular music industry revisited. *Economics Letters*, 92(1):68–74.

Giles, D. E. (2007). Survival of the hippest: life at the top of the hot 100. *Applied Economics*, 39(15):1877–1887.

Hamlen, W. A. (1991). Superstardom in popular music: Empirical evidence. *The Review of Economics and Statistics*, 73(4):729–733.

Hotelling, H. (1929). Stability in competition. *The Economic Journal*, 39(153):41–57.

International Federation of the Phonographic Industry (2024). Global music report 2024.

Irmen, A. and Thisse, J.-F. (1998). Competition in multi-characteristics spaces: Hotelling was almost right. *Journal of Economic Theory*, 78(1):76–102.

Marshall, A. (1947). *Principles of economics*. MacMillan, New York, 8th edition.

Mauch, M., MacCallum, R. M., Levy, M., and Leroi, A. M. (2015). The evolution of popular music: Usa 1960-2010. *Royal Society open science*, 2(5).

Myra, I., Kamyar, K., Lijia, W., Jienian, Y., Zhaoxia, Y., and L., K. N. (2018). Musical trends and predictability of success in contemporary songs in and out of the top charts. *Royal Society open science*, 5(5).

Neven, D. and Thisse, J.-F. (1990). On quality and variety competition. In Gabszewicz, J. J., Richard, J.-F., and Wolsey, L., editors, *Economic Decision Making: Games, Econometrics, and Optimization. Contributions in the Honour of Jacques H. Dreze*, pages 175–199. North-Holland, Amsterdam.

Pinheiro, D. L. and Dowd, T. J. (2009). All that jazz: The success of jazz musicians in three metropolitan areas. *Poetics*, 37(5):490–506. Fields in Transition, Fields in Action.

Rosen, S. (1983). The economics of superstars. *The American Scholar*, 52(4):449–460.

Spierdijk, L. and Voorneveld, M. (2009). Superstars without talent? the yule distribution controversy. *The Review of Economics and Statistics*, 91(3):648–652.

Strobl, E. A. and Tucker, C. (2000). The dynamics of chart success in the u.k. prerecorded popular music industry. *Journal of Cultural Economics*, 24:113–134.

Tabuchi, T. (1994). Two-stage two-dimensional spatial competition between two firms. *Regional Science and Urban Economics*, 24:207–227.

Terry F. Pettijohn, I. and Donald F. Sacco, J. (2009). Tough times, meaningful music, mature performers: popular billboard songs and performer preferences across social and economic conditions in the usa. *Psychology of Music*, 37(2):155–179.

Wapnick, J., Darrow, A. A., Kovacs, J., and Dalrymple, L. (1997). Effects of physical attractiveness on evaluation of vocal performance. *Journal of Research in Music Education*, 45(3):470–479.

A LLLS predictions for control variables

B Bandwidth Grid search

Optimal bandwidths are estimated by regressing each predictor in a bivariate local linear least squares (LLLS) model on the target variable *rank* for each bandwidth using 10-fold cross-validation. Therefore $N = F \times P \times B$ models are estimated with P the number of variables (regressors), B the number of bandwidths for each variable and F the number of cross-validation folds. The optimal bandwidths found with the bivariate model are not necessary optimal in the multivariate case. Because the optimal bandwidths in the multivariate model are influenced by the bandwidths of the other variables. An improved bandwidth grid search in a multivariate case should estimate the model MSE for each bandwidth combination of all variables. This would require the estimation of $N = F \times B^P$ models.

Regressor	**Optimal BW**	**Optimal BW outlier removed**
Danceability	0.1	0.1
Instrumentalness	0.35	0.6
Energy	0.05	0.05
Liveness	0.55	0.4
Valence	0.1	0.1
Acousticness	0.05	0.6
Speechiness	0.6	0.6
Artist popularity	30	50
Loudness	5	5
Duration	2.5	0.5
Tempo	15	15
Similarity 26	0.6	0.05
Similarity 52	0.6	0.05
Similarity 104	0.6	0.05

Table 3: Optimal bandwidths for each regressor

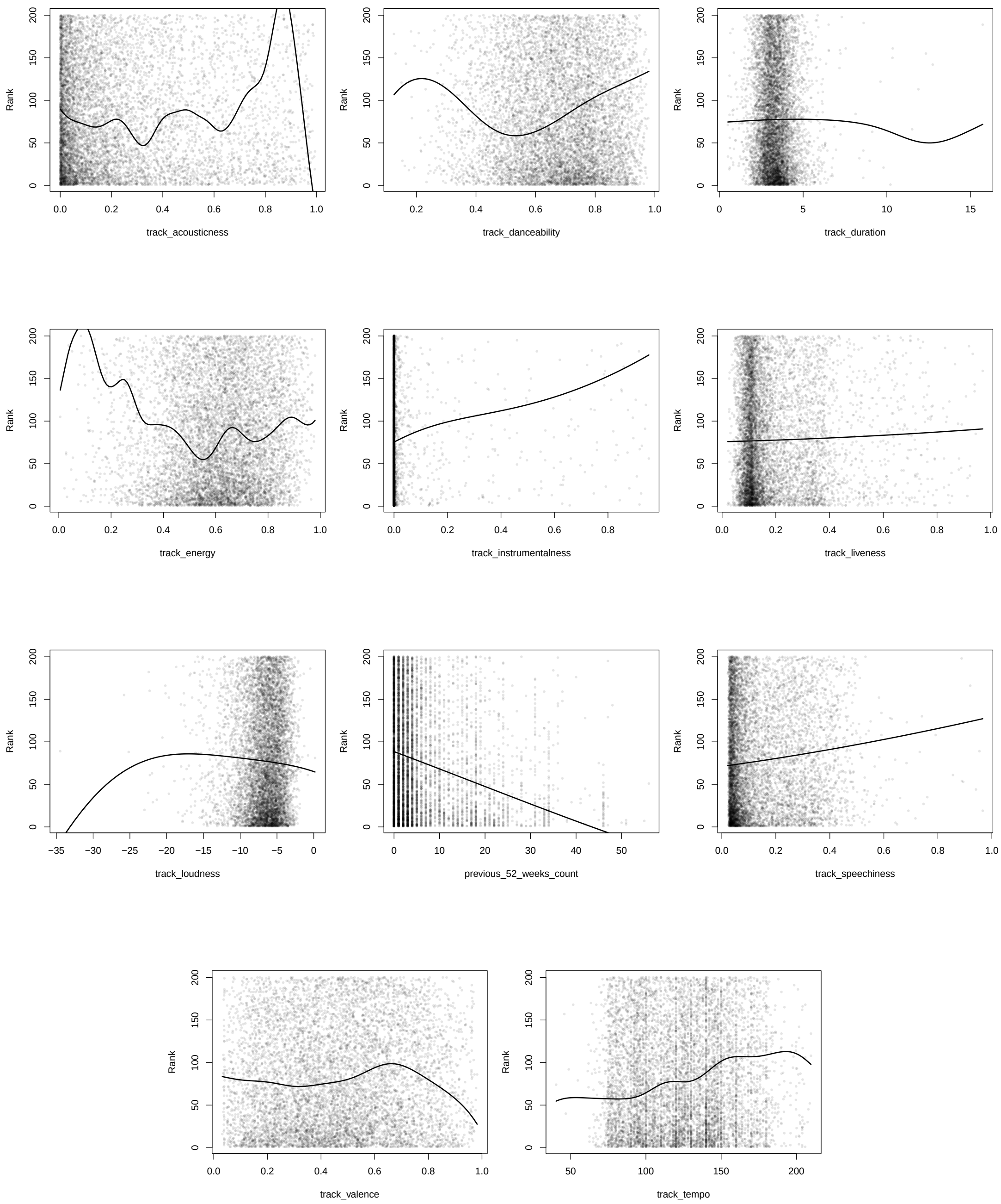

Figure 15: LLLS predictions for control variables, whole Dataset

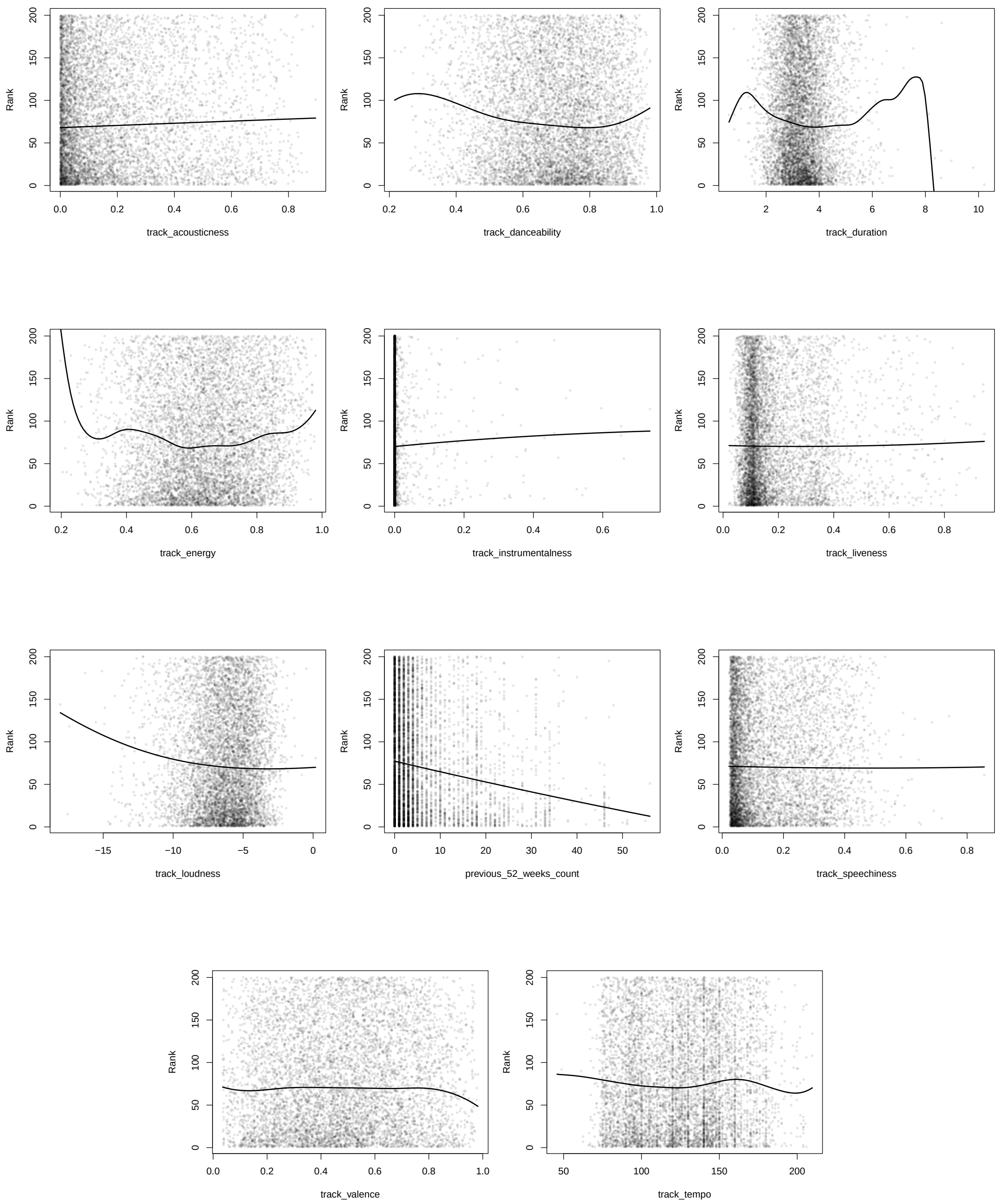

Figure 16: LLLS predictions for control variables, outlier removed

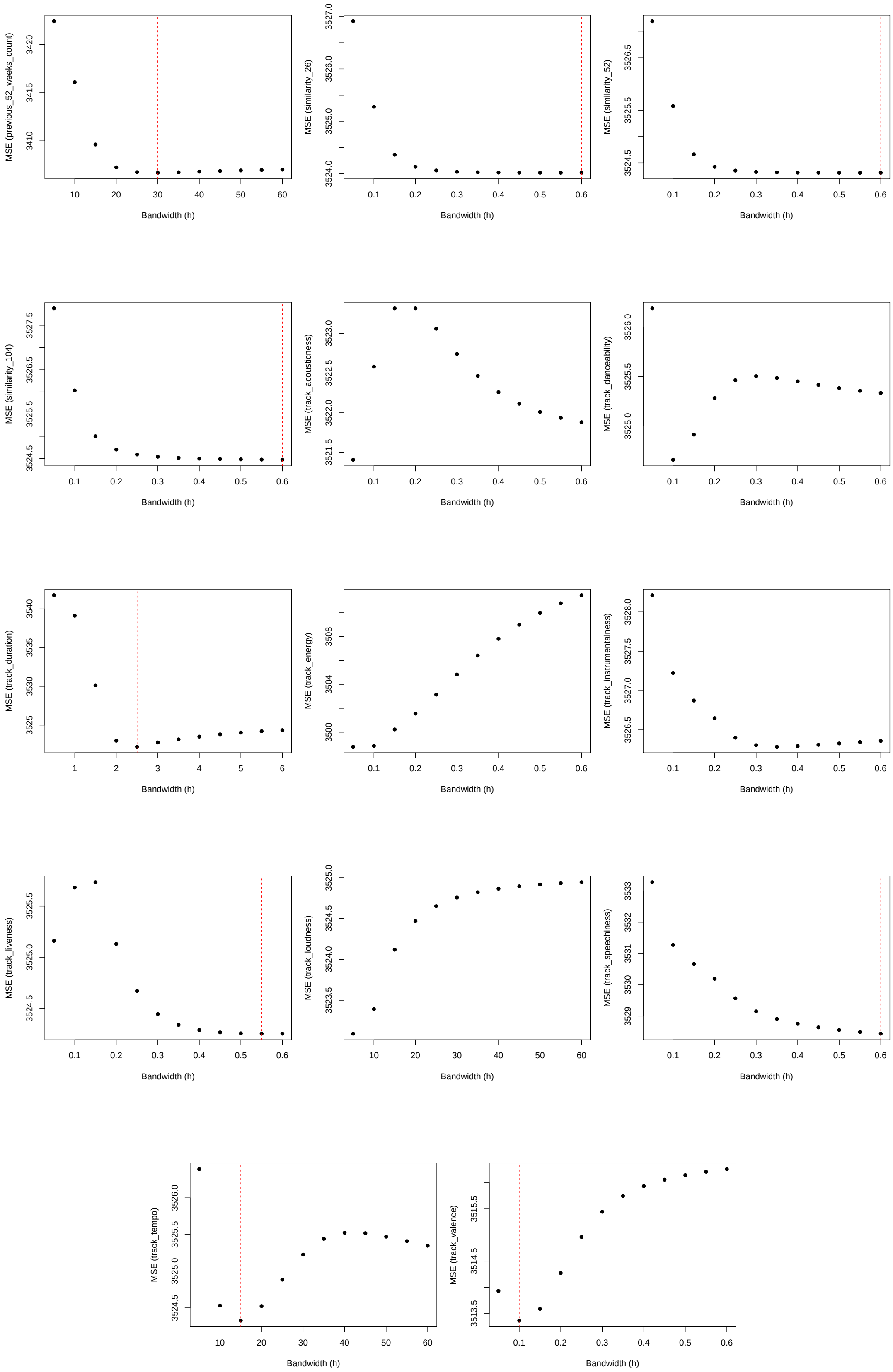

Figure 17: Bandwidth optimization for variables - whole dataset

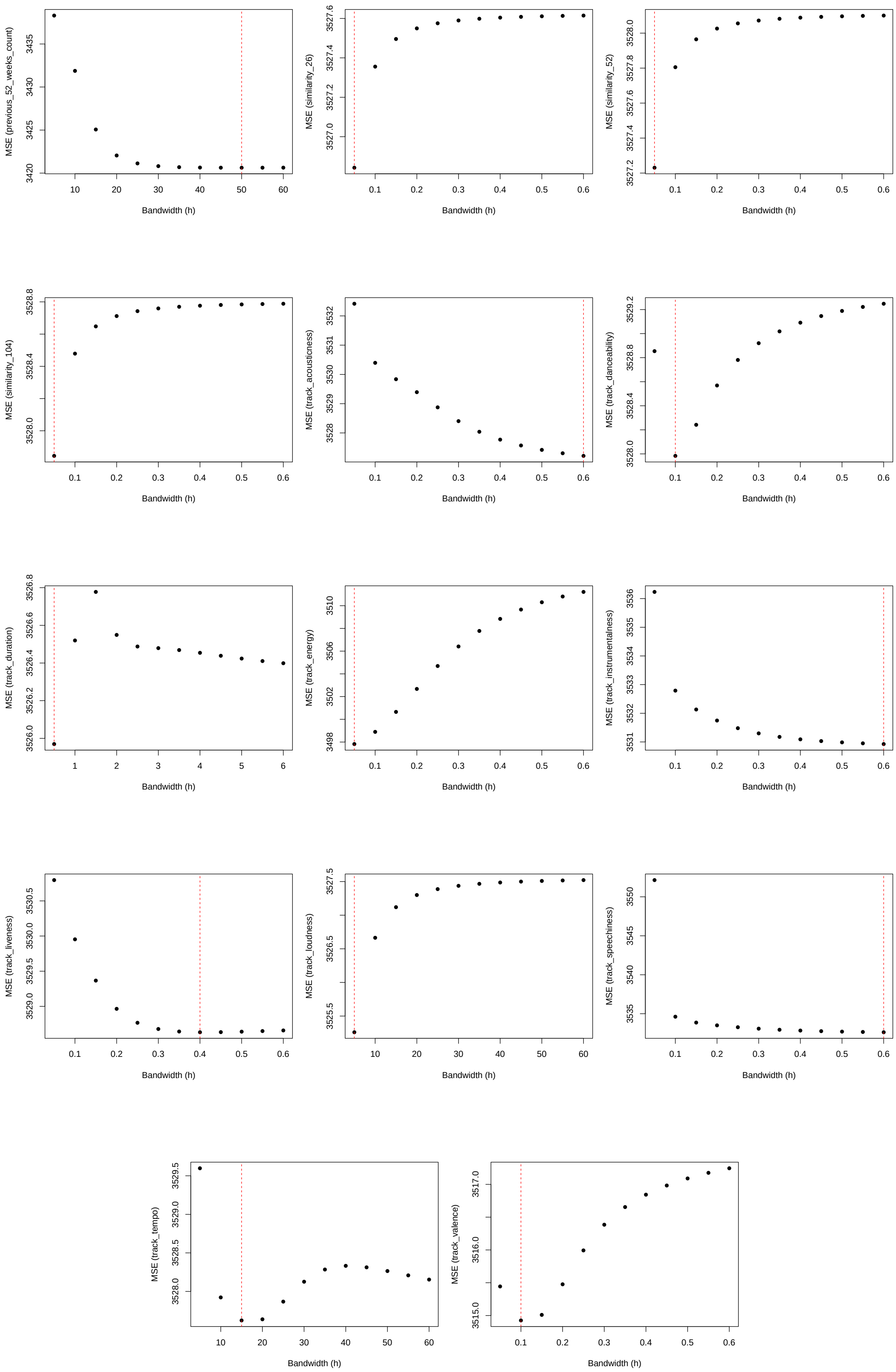

Figure 18: Bandwidth optimization for variables - outlier removed

Algorithm 1 Optimal Bandwidth Search for Regressors

```
Define bandwidths to test for each regressor
for each regressor do
    Get the bandwidth sequence for the regressor
    for each bandwidth do
        for each fold in cross-validation do
            Split data into training and testing sets
            Fit model on training set with current bandwidth
            Predict on testing set
            Compute residuals and MSE for the fold
        end for
        Compute mean MSE over all folds
    end for
    Identify optimal bandwidth for the regressor
end for
```