What Kind of Music Do You Like? A Statistical Analysis of Music Genre Popularity Over Time

Aimée M. Petitbon[1,*] and David B. Hitchcock[1]

[1]*Department of Statistics, University of South Carolina, Columbia, SC 29208, USA*

Abstract

Popular music genre preferences can be measured by consumer sales, listening habits, and critics' opinions. We analyze trends in genre preferences from 1974 through 2018 presented in annual *Billboard* Hot 100 charts and annual *Village Voice* Pazz & Jop critics' polls. We model yearly counts of appearances in these lists for eight music genres with two multinomial logit models, using various demographic, social, and industry variables as predictors. Since the counts are correlated over time, we use a partial likelihood approach to fit the models. Our models provide strong fits to the observed genre proportions and illuminate trends in the popularity of genres over the sampled years, such as the rise of country music and the decline of rock music in consumer preferences, and the rise of rap/hip-hop in popularity among both consumers and critics. We forecast the genre proportions (for consumers and critics) for 2019 using fitted multinomial probabilities constructed from forecasts of 2019 predictor values and compare our Hot 100 forecasts to observed 2019 Hot 100 proportions. We model over time the association between consumer and critics' preferences using Cramér's measure of association between nominal variables and forecast how this association might trend in the future.

Keywords *baseline-category logit; consumer preferences; Cramér's V; forecasting; multinomial logit model; partial likelihood*

1 Introduction

What kind of music do you like? People ask each other this question, both directly and indirectly, all the time. Music preference, as one of the greatest forms of self-expression and identity, says a lot about people, their values, and the types of activities they like to do. This raises the questions: What forms mass culture's music preference if we are all individuals with our unique experiences? Why do some genres reign supreme during different eras in history? We will analyze the trends in popular music genres over the past four decades and attempt to uncover what influences music genre popularity.

Past research has investigated the impact of psychology, culture, identity, and personal beliefs on music preference for individuals. At an individual level, Greenberg et al. (2015) connected music genre preference to a person's position on a scale of cognitive styles related to empathizing and systemizing. Rentfrow et al. (2011) used factor analysis to reveal five factors underlying individual-level genre preference. Andrews et al. (2020) confirmed the influence that identity, culture, and belief have on music preference, asserting that personality alone is neither the sole nor even primary influence on music popularity. Instead, personal and cultural values

---

*Corresponding author. Email: apetitbon@gmail.com.

impact music preference and music use significantly alongside personality and other factors. North and Hargreaves (2007) gave evidence that an individual's political and social beliefs influence musical preference (Andrews et al., 2020), using Hofstede's model (Hofstede, 2001) to assess whether preferences for music genres can be viewed analogously as a consumer preference (Andrews et al., 2020). Our analysis will focus on cultural, societal, and demographic factors that influence genre preference across a national culture, rather than on an individual level.

Understanding music preference helps music labels and publishing companies create a viable market for music sales. Genres, though an imperfect system, are a way to classify music and artists via a subjective measure of what certain music patterns sound like, aiming to group similar patterns of audio, instruments, and rhythms with each other. While there is a general consensus about which songs belong to certain genres, some songs may belong to multiple genres, and perceptions of genres themselves can vary depending on the listener. Music preference is a naturally inexact concept, and predefined genres tend to constrict the fluidity of music preference. Yet in the absence of any widely accepted measurement for music preference (Vlegels and Lievens, 2017), genre is commonly understood enough to be interpreted readily across wide audiences, making genres preferable to alternative categorization such as artist preference. Silver et al. (2016) suggested that genre designations are ideal because they allow musicians to construct their public images based on the generally acknowledged perceived genres and their associated cultural values. Some interesting work has been done in machine learning to automate genre classification. Jain et al. (2021) proposed a convolutional neural network algorithm to classify music genres based on timbral features. See Jain et al. (2021) for numerous relevant references on music genre classifiers.

Music sales and distribution are important for the recording industry and permit music ranking and popularity measures, such as the *Billboard* Hot 100 and *Village Voice* Pazz & Jop critics' poll, both used in our analysis. We will analyze which variables affect genre preference, examining economic and cultural factors measured on the population and measures of technological advances in recorded formats.

We measured music popularity by genre over time using two authoritative records of the annual preferences of the general population and critics' opinions. The *Billboard* Hot 100 rankings measured the popularity of genres among the general public, while *The Village Voice*'s Pazz & Jop poll rankings measured genre preferences among critics. Spivack et al. (2019) defined the *Billboard* Hot 100 rankings, which chart the top singles in the United States based on record sales, radio airtime, and online streaming (Spivack et al., 2019) as "the industry standard by which the popularity of contemporary music is measured." *Billboard* publishes Hot 100 rankings both weekly and annually; in our study, we used the annual Hot 100 data. Data from the *Billboard* Hot 100 have been analyzed previously using statistical models. Bradlow and Fader (2001) used a generalized gamma curve to model the trajectories of songs from 1993 as their weekly Hot 100 chart positions rose and declined over time. Unlike our analysis, this article did not address genre popularity over time.

The *Village Voice* Pazz & Jop poll (its name a cheeky spoonerism of *Jazz & Pop*) was a critics' opinion poll spearheaded by renowned rock critic Robert Christgau, who established the *Voice* as the "preeminent music writers' paper" (Powers, 2013), which (after an initial poll in 1971) in 1974 began annual rankings of top albums based on numerically weighted selections by a panel of critics; the poll added more ranking categories, such as singles rankings in 1979, as new recorded media developed. Christgau also incorporated more women and black writers into his critic pool, increasing the range of perspectives and genres included in rock criticism (Powers, 2013).

While the Hot 100 lists include only singles, the Pazz & Jop lists include albums throughout their timespan. Commentators such as Robins (2020) note the recent decline of the album as an artistic influence, which Hyden (2018) associates with both the decline of Rock as a genre and the newfound popularity of streaming-service playlists. While albums may no longer be as important commercially as they were in previous decades, they are still considered "cultural benchmarks" and "reliable indicator[s] of what's happening in pop culture at this very moment" (Hyden, 2018, p. 49). Furthermore, during much of the time period covered in our analysis, albums were "signposts for when cultural movements rose to prominence and . . . crested" (Hyden, 2018, p. 50), making their inclusion invaluable in our analysis of genre trends over time.

In the early years of the poll, a perspective later derided as Rockism, which championed the preservation of values inherent to Rock — the iconic genre — dominated critics' opinions. Sanneh (2004), in an influential article, espoused a counterpart movement that became known as Poptimism, which held that stars of genres such as Pop, Rap, and R&B were as worthy of critical approbation as artists in the theretofore more "authentic" genres, like Rock and Folk (Harvilla, 2017). This movement became de rigeur in critical circles throughout the 2000s and 2010s, as artists such as Justin Timberlake, Beyoncé, Kanye West, and Taylor Swift converted popular fame into critical acclaim. Richards (2015) and Harvilla (2017) noted some degree of backlash against Poptimism, suggesting that it led to the celebration of shallower art. However, Hyden (2018) and critic Jon Caramanica (Harvilla, 2017) attribute the end of Rock's dominance and Rockism to the waning influence of rock radio and the explosion of streaming formats, whose algorithm-based playlists push listeners toward famous pop stars. The decline of the Rock genre in consumer and critical appeal will be one of the themes we will investigate in our analyses.

2   Data and Variables

We collected the annual *Billboard* Hot 100 Singles data for the years 1974 through 2018 (as well as 2019 data which we used to validate model forecasts). For the Pazz & Jop poll, we recorded the top 20 albums and top 20 singles (including tied rankings) for years 1974 through 2018. Since the Pazz & Jop poll was produced each year from 1974 to 2018, we defined 1974 to 2018 as the timeframe of our data set, a time period which encompasses much of the history of rock music and serious critical coverage of popular music.

For each year, we counted the number of appearances in the Hot 100 ranking and the Pazz & Jop poll for each of eight genres (ignoring the artist and song title once the genre was determined since our interest was only in tracking genre popularity). The Pazz & Jop singles and albums counts were combined by year, resulting in genre counts that were the total number of appearances for each genre in the top 20 singles and top 20 albums for each year. We used only the top 20 for consistency across years since the early polls only listed 20 entries for albums and for singles, though no singles were ranked in the Pazz & Jop poll until 1979, so the first five years had lower total genre counts than the remaining years.

The genre assignments were adapted from an impressive project by The DataFace (Beckwith, 2016), a data visualization agency, which used the Spotify API and web searches to categorize the songs into genres. While about 79% of genres for songs in their data set came from the Spotify API, Beckwith (2016) used a two-step process to recategorize the genres according to their study's specifications of sixteen genres. Beckwith acknowledged their analysis and assignment methodology was imperfect since Spotify assigns multiple genres at the artist level, not by song, and that their study simplified this by giving each artist the same genre regardless of song

or album. Note that a small number of genres assigned by The DataFace were not appropriate based on consensus and authoritative judgment, which reflected their explanation of the blanket artist/genre assignment approach. In certain cases, artists clearly changed genres in their careers; for example, Taylor Swift switched from Country to Pop and Bruce Springsteen moved back and forth between Folk and Rock during his career. With our discretion, for these and other artists, we manually adjusted the genre assignments in the Hot 100 and Pazz & Jop data sets to reflect such genre shifts and to ensure consistency of results across both data sets. This genre review was necessarily subjective, but almost certainly resulted in more accurate genre designations in the cases where corrections were made. We assigned genres for all Pazz & Jop entries and for the Hot 100 entries after mid-2016, maintaining consistency with the DataFace designations where possible. Some singles and albums were not listed in the DataFace data and required web searches to assign an appropriate genre.

Note that the original DataFace source categorized songs into sixteen unique genres. To reduce the complexity of the multinomial response variable, we collapsed the sixteen genres into eight by combining similar genres and creating an "Other" category. Similar genres (R&B and Soul; Pop Standards and Pop; Rock & Roll, Rock, Punk, and Metal) were combined. In the "Other" category, we placed genres which had lower counts, such as Latin, Blues, Jazz, and Ska/Reggae/Dancehall. Most of these genres in the "Other" category shared the unifying characteristic of originating within a specific American ethnic community (or being imported from a foreign culture) and retaining a following that was relatively limited to a subculture rather than ever attaining hegemony in American culture.

Tables 1 and 2 show the counts for the eight genres for a selection of the sampled years for the Hot 100 data and Pazz & Jop data, respectively. For example, the first row in Table 1 shows that during 1974, the total number of Country songs in the Hot 100 list was 10, the number of Folk songs was 11, and so on. The first row of Table 2 indicates that during 1974, the number of Country albums in the top 20 was 2, the number of Folk albums was 4, and so on. (Recall that the first five years of the Pazz & Jop list included only albums, not singles.) Genre count tables for all years (1974-2018) for the Hot 100 and Pazz & Jop data, as well as raw data files with singles and albums and their genre assignments, can be found in the Supplementary Material.

We considered candidate predictors (each measured annually) that were economic, demographic, and technological factors that may influence music genre popularity over time in the United States from 1974 to 2018. Note that we considered lag-1 versions of some predictors as well, in case the predictor's effect on the genre counts is delayed by one year.

While our models are exploratory in nature, there are several related studies that have investigated the effects on music consumption and genre preference of variables similar to those we consider in our models. Liu et al. (2018), one of the few articles other than ours to assess music genre tastes at the country level, measured genre preferences as well as album and artist preferences across countries. Similar to us, Liu et al. (2018) used GDP per capita as a proxy for economic status although they did not find a significant economic effect on genre preferences at the country level. The factors in their models were "based on several prominent theories in sociology and psychology (Bourdieu's class theory, Social identity theory and Hofstede's cultural dimension theory)," (Liu et al., 2018, p. 5). Unlike ours, their study compared different countries cross-sectionally rather than one country's genre preferences over time. Woolhouse and Bansal (2013) also analyzed country-level economic variables, such as unemployment rate, which had a significant effect on music consumption. The analysis of Bryson (1996) related political tolerance to music genre tolerance similarly to the political ideology factor in our models, although their analysis was on the individual level rather than the country level.

Table 1: Genre Counts (1974-1985, 2017-2018) for Annual *Billboard* Hot 100.

					Genre Counts for Hot 100 data, Selected Years			
Year	Country	Folk	Elec	Pop	Rap/Hip-Hop	Rock	RBSoul	Other
1974	10	11	1	11	0	30	34	3
1975	5	8	1	7	0	46	30	3
1976	3	2	1	9	0	44	39	2
1977	3	5	0	17	0	49	24	2
1978	5	6	2	8	0	49	28	2
1979	6	3	5	8	0	43	34	1
1980	10	6	0	8	0	51	24	1
1981	16	6	3	9	0	48	15	3
1982	9	4	2	4	0	64	16	1
1983	2	1	1	13	0	67	13	3
1984	2	0	2	16	0	56	23	1
1985	0	0	3	15	0	57	24	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
2017	6	0	9	37	38	2	6	2
2018	9	0	3	30	43	2	8	5

Table 2: Genre Counts (1974-1985, 2017-2018) for Pazz & Jop. Counts for Pazz & Jop for 1974-1978 include only albums, not singles.

					Genre Counts for Pazz & Jop data, Selected Years			
Year	Country	Folk	Elec	Pop	Rap/Hip-Hop	Rock	RBSoul	Other
1974	2	4	0	0	0	12	2	0
1975	2	3	0	0	0	12	1	2
1976	0	4	0	0	0	12	3	1
1977	0	2	0	0	0	16	1	1
1978	1	0	0	0	0	18	0	1
1979	0	2	0	5	0	27	6	1
1980	0	2	0	2	1	28	7	0
1981	1	0	3	2	1	22	9	2
1982	0	2	2	2	2	21	7	4
1983	0	2	1	4	4	24	5	2
1984	0	0	2	4	3	24	5	2
1985	0	2	0	5	1	26	7	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
2017	0	3	2	12	10	8	5	0
2018	7	1	1	10	8	10	3	2

Inflation rate (INFL) as a percentage, calculated from the Consumer Price Index, was obtained for each year from the Bureau of Labor Statistics (Coinnews Media Group LLC, 2021). We

gathered annual unemployment rates (UNEMP) by averaging the monthly rates (Macrotrends, 2021b) for each year. We collected annual growth rates of the United States' gross domestic product (GDP) from (Macrotrends, 2021a).

To measure the political ideology of the United States population over time, we used the results of a political survey question asked as part of the General Social Survey (GSS Data Explorer, 2021). A random sample of the U.S. population was asked to rate themselves as liberal or conservative on a seven-point scale (1=most liberal, 7=most conservative). For each year, we used the mean response value, which we labeled POLIT, to encapsulate the average U.S. ideological position for that year. The question was asked annually from 1974 through 1993 (except 1979, 1981, and 1992) and every even year from 1994 to 2018; we used linear interpolation to impute values for the missing years, yielding a yearly time series from 1974 to 2018.

For inflation-adjusted sales revenue (SALES), given in millions of U.S. dollars (USD) by the Recording Industry Association of America (Recording Industry Association of America, 2021b), we used the `adjust_for_inflation` function in the `priceR` package (Condylios, 2020) in R to adjust the revenue values for inflation into 2018 USD to render values comparable across years, which is essential to model the true effect that sales revenue has on genre popularity.

For the proportion of physical formats (PHYS), we collapsed the listed formats (Recording Industry Association of America, 2021a) from the RIAA sales data into a binary variable: (physical/streaming). For example, the sales of "LP/EP", "Cassette", "8-Track", "Cassette Single", "Vinyl Single", "CD", "CD Single", "DVD Audio", "Kiosk", and "Other Tapes" were physical, and all other formats were streaming. In recent years, streaming music has become increasingly prevalent, as smartphones dominate the streaming industry (Pedrero-Esteban et al., 2019) thanks to attractive visuals, interactive apps, and services that access streaming immediately, such as Spotify, YouTube, and Apple Music. All these services benefit from the advanced technology in various smart devices that previous formats could never capture. Including the PHYS variable in the model helped determine whether the prevalence of format types affects genre popularity.

In the Supplementary Material, we provide a scatterplot matrix of our candidate predictor variables for each model. With few exceptions such as the proportion physical variable, most predictors display relatively little pairwise association, although all are measured over time. However, the partial likelihood model will still necessitate careful (conditional) interpretations of the predictors' effects.

3   Approach and Methodology

In this section, we outline our methods of analyzing the data from the *Billboard* Hot 100 on consumer genre preferences and the *Village Voice* Pazz & Jop poll on critics' genre preferences. These provide a wide scope of genre preference data to analyze the changes in popular music from 1974 through 2018 — almost the entire timespan of the era of rock music and popular music criticism. The over forty-year time span offers enough data for proper time series analysis and the use of two separate multinomial logit regression models, one for each data set. All analyses were done in R (R Core Team, 2020). We used the `vglm` function in the `VGAM` package (Yee, 2021) to build the multinomial logit models. As a generalized linear model, the multinomial (baseline-category) logit model is appropriate to fit the genre count response variables (Agresti, 2015) which are nominal categorical responses. Since at least during the initial part of our timeframe, Rock was the dominant category in both the Hot 100 and Pazz & Jop data, we chose Rock

as the baseline category, and this designation allowed us to interpret other genres' changes in popularity over time relative to Rock.

For dependent time series data, autocorrelations among predictors, responses, or both need to be accounted for to ensure a good model fit with accurate and interpretable results. The time series genre count variables in this analysis, combined with the time series for each covariate included in the model, follow an unknown joint distribution. The partial likelihood approach uses one piece of the factored joint density as a simpler type of likelihood function that does not lose much information under typical conditions. Fokianos and Kedem (2004) showed that using partial likelihood inference with generalized linear models accounts for autocorrelation among responses and covariates, which our data set contains. Their approach allows for a simple implementation of the partial likelihood estimation that can be accomplished with statistical software packages such as R, since their partial likelihood score equation matches that of a model with independent data (Fokianos and Kedem, 2004). The partial likelihood method relies only on the conditional distribution of the current response and relates the conditional mean $\mu_t = E[Y_t|\mathcal{F}_{t-1}]$ to the data through the time of actual observation (Fokianos and Kedem, 2004), where $\mathcal{F}_{t-1}$ is "the σ-field generated by past values of the response series and past and possibly present values (when known) of the covariates" (Fokianos and Kedem, 2004, p. 176). Therefore, we can proceed with our analysis despite not knowing the full joint distribution for the multiple time series. In practice, following the approach of Fokianos and Kedem (2004), we condition on $\mathcal{F}_{t-1}$ by including the lagged response variables as covariates to account for autocorrelations between each year. While the data collected ranges from 1974 to 2018, the response data analyzed is from 1975 to 2018 to allow us to include the lagged responses from 1974 to 2017 as predictors. As mentioned, due to the time series nature of the predictors, we also considered lagged versions of the predictors.

Finally, while generalized linear mixed models (Breslow and Clayton, 1993) and generalized estimating equations (Liang and Zeger, 1986) are commonly used for generalized linear models with time-dependent responses, such methods are ideal for panel data with repeated measurements on numerous individuals. Our data sets, however, can be viewed as having multiple variables measured repeatedly over time on a single individual (the United States), making the multivariate time series framework of Fokianos and Kedem (2004) the most appropriate approach.

4 Multinomial Logit Model Building

Our strategy for model selection is as follows: We consider all our demographic, economic, and music-industry variables (UNEMP, INFL, POLIT, GDP, SALES, and PHYS) as candidate predictors, along with lag-1 versions of UNEMP, INFL, POLIT, GDP, and SALES in case the effect of any of these variables on genre popularity is delayed. We fit multinomial logit models with all possible subsets of the set of candidate predictors. We rank the candidate models based on BIC (Schwarz, 1978) and choose our best model, considering both BIC and AIC (Akaike, 1973) holistically. Our use of BIC and AIC for model selection follows the example of Fokianos and Kedem (2004) in their foundational article on partial likelihood, as well as Hosseini et al. (2012) who used BIC and AIC in a similar generalized linear model fit using partial likelihood. At this point in the variable selection, we do not include time itself as a candidate predictor because we do not want time to push meaningful predictors out of the model and serve as a surrogate for them, but we consider later whether adding a time trend to the final model

improves the fit. Once the "best" set of predictors is chosen, we add lag-1 versions of our genre counts as predictors in the model. This improves the fit and predictive ability of the model, but the major reason for including lagged dependent variables (LDVs) as predictors is to implement the partial likelihood approach of Fokianos and Kedem (2004) and have the model account for the correlation over time of the genre counts. Finally, we consider including linear, quadratic, or cubic time trends, judging based on BIC and AIC whether the inclusion of time improves the final model.

4.1 Model Selection and Fit for Hot 100 Data

Once we fit models using all possible subsets of our candidate predictors, we chose the model with the lowest BIC value (see Table 3) which had the following predictors: $UNEMP_t$, $INFL_{t-1}$, $POLIT_{t-1}$, GDP_t, $PHYS_t$, $SALES_t$, and $SALES_{t-1}$, where the subscript $t-1$ denotes the lag-1 version of the predictor. Our chosen model minimized BIC and had the sixth lowest AIC value, but models with lower AIC had substantially worse BIC values. We then included the lagged dependent variables as predictors following the partial likelihood approach. Lastly, we considered whether adding time trends would improve model fit. We sequentially added linear, quadratic, and cubic time terms; the model with the cubic time trend had the lowest AIC while the model with time omitted had the lowest BIC. The in-sample fits for both models were nearly identical. Ultimately, we chose the more parsimonious model without time as our best model. In Section 5, we will compare forecasts for future Hot 100 genre probabilities for the two models.

Averaged political ideology and inflation rate entered the model in lagged versions because these predictors required a year for their respective effects on genre popularity to take effect. Intuitively, this is reasonable considering these are economic and demographic changes that would not create an immediate impact. Both the lagged and unlagged versions of inflation-adjusted sales were important in the model.

We analyzed residuals of Model 1 to further assess model fit. See the Supplementary Material for all the following plots. For generalized linear models, the magnitudes of the raw response residuals are not particularly meaningful, although looking at the general pattern of the residuals plotted against time may be instructive. More useful are the Pearson residuals. The scattered and mainly nonpatterned residuals across the genres for our chosen model resemble white noise for most of the genres and suggest that the partial likelihood model choice with the lagged dependent variables as predictors has accounted for the autocorrelations in the responses. Note that the Pearson residuals are defined relative to the baseline category, so plots for only the seven non-baseline genres (excluding Rock) are shown. Some runs of near-zero residuals, such as with Rap/Hip-Hop, occurred during periods when the counts for that genre were extremely low. The autocorrelation function plots of the residuals show that most autocorrelations are within the two-standard-error bounds, indicating that the residuals are not significantly correlated across time for each genre, respectively. The 4×2 scatterplot depicts the Pearson residuals plotted against their lag-1 analogs and further confirms the lack of lag-1 autocorrelation. The bottom right panel of this figure plots the Pearson residuals against the model's fitted values and establishes the overall goodness of fit of the model.

It may be instructive to examine some of the estimated coefficients of the model to interpret the effects of certain covariates on the popularity of specific genres relative to Rock. This could help explain which factors drive the changes in genre popularity over time. Since the partial likelihood model is a conditional one, the estimated coefficients and associated odds ratios must be interpreted conditionally (Fokianos and Kedem, 2004). As an example, we interpret an effect

Table 3: *Billboard* Hot 100 Multinomial Logit Model Selection Results for Selected Candidate Models Before Inclusion of Lagged Dependent Variables and Consideration of Time Terms.

Model	Predictors	AIC	BIC
Model 1	$\text{UNEMP}_t + \text{POLIT}_{t-1} + \text{INFL}_{t-1} + \text{GDP}_t +$ $\text{PHYS}_t + \text{SALES}_t + \text{SALES}_{t-1}$	1619.2	1719.2
Model 2	$\text{UNEMP}_{t-t} + \text{POLIT}_t + \text{POLIT}_{t-1} + \text{INFL}_{t-1} +$ $\text{PHYS}_t + \text{SALES}_t + \text{SALES}_{t-1}$	1619.5	1719.4
Model 3	$\text{UNEMP}_{t-1} + \text{POLIT}_{t-1} + \text{INFL}_{t-1} +$ $\text{PHYS}_t + \text{SALES}_t + \text{SALES}_{t-1}$	1632.2	1719.6
Model 4	$\text{UNEMP}_{t-1} + \text{POLIT}_{t-1} + \text{INFL}_{t-1} + \text{GDP}_t +$ $\text{PHYS}_t + \text{SALES}_t + \text{SALES}_{t-1}$	1624.3	1724.2
Model 5	$\text{UNEMP}_t + \text{POLIT}_{t-1} + \text{INFL}_{t-1} + \text{GDP}_t +$ $\text{GDP}_{t-1} + \text{PHYS}_t + \text{SALES}_t + \text{SALES}_{t-1}$	1612.8	1725.2
Model 6	$\text{UNEMP}_t + \text{POLIT}_{t-1} + \text{INFL}_{t-1} + \text{GDP}_t +$ $\text{PHYS}_t + \text{SALES}_{t-1}$	1637.9	1725.4

associated with the unemployment rate predictor: For a fixed year, conditional on the current and past values of the predictors and the past values of the response, the odds of a slot in the Hot 100 list being Country, relative to Rock, are $\exp(0.22) = 1.25$ times (i.e., a 25% increase) what they would be if that year's unemployment rate were 1% lower than it is. This implies that a higher overall unemployment rate is associated with higher odds of Country relative to Rock. As another example, interpreting one of the effects associated with the lagged political ideology predictor: For a fixed year, conditional on the current and past values of the predictors and the past values of the response, the odds of a slot in the Hot 100 list being Pop, relative to Rock, are $\exp(0.44) = 1.55$ times larger than they would be if the previous year's average political ideology score was 0.1 lower than it is for that year. This implies that a higher, or more conservative, lagged political score is associated with higher odds of Pop relative to Rock. Table 4 lists selected odds ratio estimates. Note that the odds ratio for Country relative to Rock (0.64), which would imply that Country's popularity rises as the national political ideology grows more liberal, appears counterintuitive to popular conceptions of conservative Country music enthusiasts. Possibly, this effect could reflect that Country music fans might buy more music as a reaction against a more liberal political climate. Because our model has eight response categories and numerous predictors, the number of coefficients is quite large. In the Supplementary Material, we provide a list of all the estimated coefficients of the model, which can be used for similar interpretations.

Table 4: Estimated odds ratios for each genre relative to Rock corresponding to effects of unemployment rate predictor and to (a 0.1 increase in) the lagged political ideology predictor.

Predictor	Country	Folk	Elec.	Pop	Rap/ Hip-Hop	RBSoul	Other
Unemployment Rate	1.25	1.44	1.34	1.03	1.00	1.02	0.78
Lagged Political Ideology	0.64	2.80	1.18	1.55	1.46	1.19	0.43

4.2 In-Sample Model-Predicted Genre Probabilities for Hot 100

The almost exact overlay of observed genre proportions and model-predicted probabilities for each of the eight genres in Figure 1 demonstrates how accurate the chosen model was in fitting the Hot 100 data. In the few places where the model-predicted probabilities were not perfectly overlain, the model tended to smooth the rising and falling trends over time present across the various genres.

After dominating the 1980s thanks to broadly popular artists like Bruce Springsteen, Duran Duran, and The Police, Rock steadily declined in the late 1980s through 2000s, opening the door for the steady rise of Rap/Hip-Hop and the bursts of popularity for Pop in the early 1990s and 2010s. Rap/Hip-Hop experienced surges in the early 1990s from artists like MC Hammer, Dr. Dre, and Salt-N-Pepa, in the early 2000s from emerging talent such as Jay-Z, Nelly, and Eminem, and in the mid-2010s driven by Drake. Pop's sporadic rises featured multimedia stars such as Madonna, Michael Jackson, George Michael, and Whitney Houston in the late 1980s and then Katy Perry, Rihanna, and Chris Brown in the Poptimism era of the early 2000s. After a short-
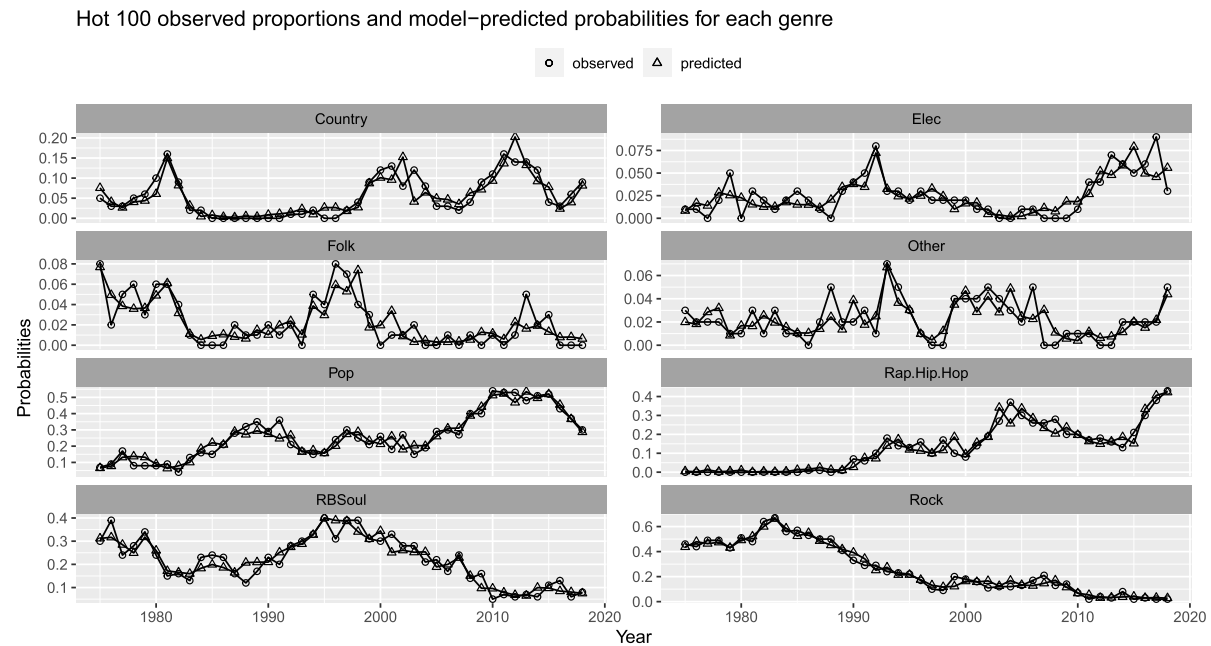

Figure 1: Observed proportions (circles) and model-predicted probabilities (triangles) from 1975-2018 for each of the eight genres, for the Hot 100 data.

lived uptick in the early 1980s led by Kenny Rogers, Country notably rose in popularity from 1999 through 2003 thanks to mainstream success of artists like Lonestar, the Dixie Chicks, Kenny Chesney, Tim McGraw, and Shania Twain and peaked again from 2008 through 2013 led by Taylor Swift during the Country phase of her career. R&B/Soul saw a relative decline until the 1990s when, due to strong sales from artists like TLC, Boyz II Men, and Janet Jackson, the category rose back up in popularity to match its high point from the 1970s before declining in popularity through 2018. Finally, Electronic rose in popularity in the 2010s driven by the success of David Guetta.

4.3 Model Selection and Fit for Pazz & Jop Data

We now present our analysis of the Pazz & Jop data, which reflected critics' genre preferences across time. After fitting models using all possible subsets of our candidate predictors, the model with by far the lowest BIC value, although other models had lower AIC values (see Table 5), had the following predictors: INFL_{t-1} and PHYS_t.

   We again included the lagged dependent variables as predictors and checked whether including time trends would improve the model fit. In this case, a model with a quadratic time trend reduced the AIC and yielded more convincing in-sample fits for certain genres like Country, so our final model included this quadratic time trend.

   From an interpretation standpoint, our chosen Pazz & Jop model is simpler than the Hot 100 model, which included numerous demographic predictors. Since professional critics' opinions are likely based on deep-seated philosophies about music quality, it seems reasonable that critical preferences are less malleable and less affected by societal factors than consumer preferences. Furthermore, it is logical that one important factor that appears to influence critics' genre preferences reflects arguably the most consequential development in the commercial music industry in this timeframe: the introduction of streaming formats as an alternative to traditional physical

Table 5: *Village Voice* Pazz & Jop Multinomial Logit Model Selection Results for Selected Candidate Models.

Model	Predictors	AIC	BIC
Model 1	$\text{INFL}_{t-1} + \text{PHYS}_t$	1155.0	1192.5
Model 2	$\text{INFL}_{t-1} + \text{PHYS}_t + \text{SALES}_{t-1}$	1151.5	1201.4
Model 3	$\text{POLIT}_{t-1} + \text{INFL}_{t-1} + \text{PHYS}_t$	1152.2	1202.1
Model 4	$\text{INFL}_t + \text{PHYS}_t$	1165.2	1202.7
Model 5	$\text{INFL}_{t-1} + \text{PHYS}_t + \text{SALES}_t$	1153.2	1203.1
Model 6	$\text{UNEMP}_{t-1} + \text{INFL}_{t-1} + \text{PHYS}_t$	1153.6	1203.5

Table 6: Estimated odds ratios for each genre relative to Rock corresponding to the effect of the lagged inflation rate predictor and to (a 0.1 increase in) the "proportion of physical formats" predictor.

Predictor	Country	Folk	Elec.	Pop	Rap/ Hip-Hop	RBSoul	Other
Lagged Inflation Rate	1.15	0.98	1.22	1.08	0.90	1.02	0.91
Proportion of Physical Formats	1.87	0.84	0.70	0.78	0.85	0.49	1.42

formats (Hiller and Walter, 2017). One might speculate that the availability of music in streaming formats allowed critics easier access to more types of music than in the past, and potentially allowed for more critical recognition of different genres. Figures in the Supplementary Material exhibit residual analysis of this model to further assess model fit. The Pearson residuals mostly resembled white noise and their autocorrelations were within the two-standard-error bounds, indicating a lack of serial dependence. The 4×2 scatterplots revealed no issues with model fit.

We now interpret effects of covariates in the Pazz & Jop model, similar to those provided with the Hot 100 model. Table 6 lists the estimated odds ratios for each genre. For example, interpreting the effect of the covariate measuring the proportion of sales in physical formats: For a fixed year, conditional on the current and past values of the predictors and the past values of the response, the odds of a slot in the Pazz & Jop list being Country, relative to Rock, are $\exp(0.6248) = 1.87$ times (i.e., a 87% increase) what they would be if that year's proportion of sales in physical format were 0.1 lower than they are for that year. (Note that because this covariate is a proportion, the hypothetical change we examine is 0.1 units, which is a more meaningful change for this variable.) This implies that a higher overall prevalence of physical formats is associated with higher odds of Country relative to Rock. As another example: For a fixed year, conditional on the current and past values of the predictors and the past values of the response, the odds of a slot in the Pazz & Jop list being Rap/Hip-Hop, relative to Rock, are $\exp(-0.1646) = 0.85$ times (i.e., a 15% decrease) what they would be if that year's proportion of sales in physical format were 0.1 lower than they are for that year. This implies that a higher overall prevalence of physical formats is associated with lower odds of Rap/Hip-Hop relative to Rock. The Supplementary Material includes all the estimated coefficients of the model, which can be used for similar interpretations.

4.4 In-Sample Model-Predicted Genre Probabilities for Pazz & Jop

As shown in Figure 2, the Pazz & Jop model had less of an exact overlay of observed genre proportions and model-predicted probabilities for the eight genres possibly due to the greater variability of critics' opinions over time, but the model captured a keen likeness of the true trends through the 40 years and smoothed some of the noise in the yearly counts. As with the Hot 100, we see the rise of Rap/Hip-Hop from 1986 to 1992, exemplified by the critically praised acts Public Enemy and the Beastie Boys. Rock saw a gradual decline — although less severe with critics than with consumers — that was only temporarily halted by the alternative rock movement in the early 1990s, spearheaded by bands like R.E.M. and Nirvana. Critics consistently placed more importance on Folk as a genre than did consumers. Pop became more acceptable in the eyes of critics with a steady rise from the mid-1990s to the mid-2010s, reflecting the end of
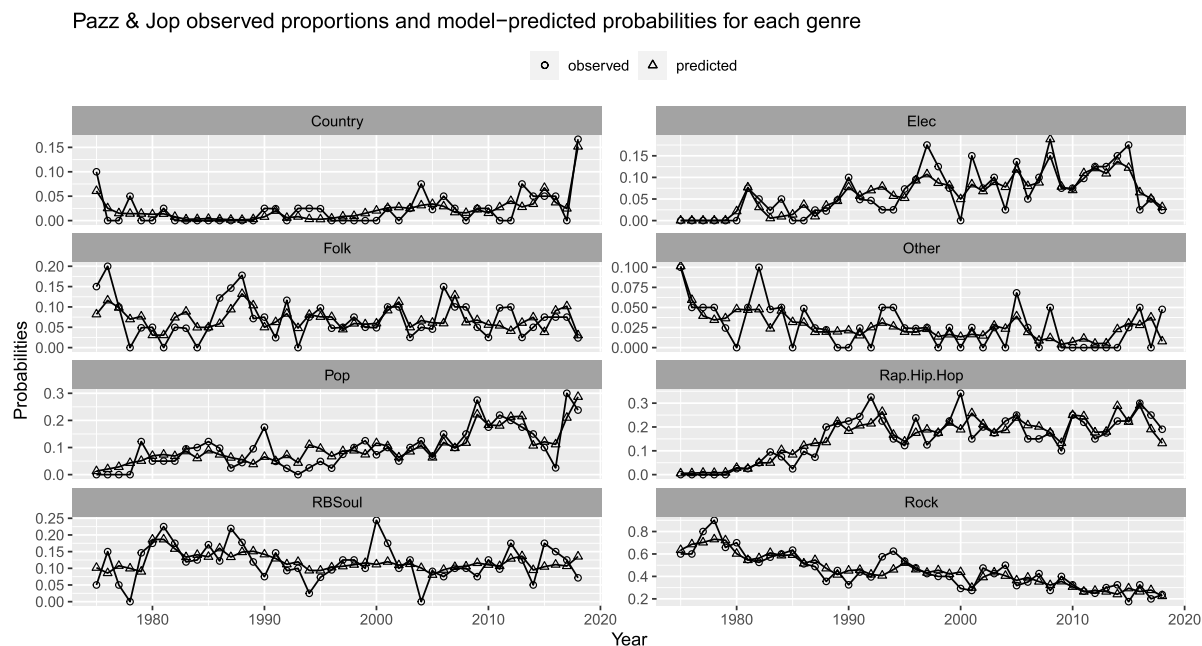
Figure 2: Observed proportions (circles) and model-predicted probabilities (triangles) from 1975-2018 for each of the eight genres, for the Pazz & Jop data.

Rockism and the increase in broader critical acceptance of a wider class of genres. Pop, which in the mid-2010s experienced a downturn in critical esteem, perhaps due to the Poptimism backlash noted by Harvilla (2017), saw a resurgence after 2016. Electronic music similarly gained favor among critics in the 2010s. Despite not being as favored among critics as among consumers, Country enjoyed a gradual rise in critics' appreciation throughout the early 2000s before falling off the critics' radar in 2017 and rebounding in 2018 thanks to Kacey Musgraves' acclaimed *Golden Hour* album and accompanying singles.

5   Forecasting Future Genre Proportions

We now present an approach for forecasting future genre proportions using our model. Since our model depends on lag-1 genre counts as predictors, we focus on forecasts one year into the future (say, at time $n + 1$, which is the year 2019 in our case), since genre counts at time n will have been observed. Unlagged values of the predictors at time $n + 1$ will not be available, so for these, we fit an ARIMA model to each series of predictor values through 2018 using the `auto.arima` function in the `forecast` package (Hyndman et al., 2020) of R – choosing the model that minimizes AIC – and plug the forecasted values in place of the predictors in the fitted model equation. To constrain the forecasted value of the proportion of physical sales to be between 0 and 1, we simulate from an ARIMA model fitted to the logit-transformed series and then back-transform that simulated value. (For proportions of exactly 1 in the series, we follow the recommendation of Warton and Hui (2011) and add to the numerator and denominator of the logit function the quantity ϵ = the minimum nonzero value of 1 minus each observed value.)

To account for the uncertainty in these future predictor values, we simulate the next year's values 200 times from each ARIMA fit. To account for the sampling variability in the estimated
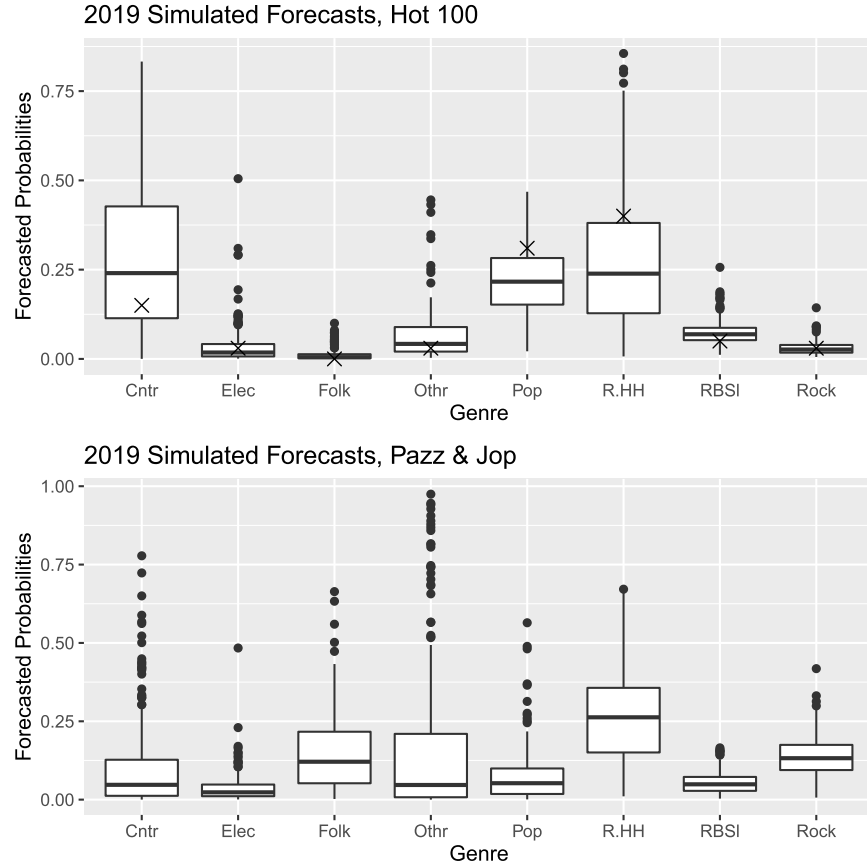
Figure 3: Top: Simulated Multinomial Logit Forecast Boxplots for Hot 100 data for 2019. Actual genre proportions for 2019 Hot 100 rankings are displayed with ×. Bottom: Simulated Multinomial Logit Forecast Boxplots for Pazz & Jop data for 2019.

coefficients, we generate 200 sets of coefficients $\boldsymbol{\beta}^{[1]}, \ldots, \boldsymbol{\beta}^{[200]}$ from a multivariate normal distribution with mean vector equaling the estimated coefficients $\hat{\boldsymbol{\beta}}$, since Fokianos and Kedem (2004) show the maximum partial likelihood estimates are consistent for $\boldsymbol{\beta}$ and are asymptotically multivariate normal. The covariance matrix of the coefficient estimates was estimated using a jackknife approach because the standard covariance formulas may not apply in the partial likelihood case. The output from plugging these and the relevant (simulated for some components) covariate vectors, say, $\mathbf{X}_i^{[j]}$ into the model-predicted probability equation

$$\hat{\pi}_i^{[j]} = \frac{\exp(\mathbf{X}_i^{[j]}\boldsymbol{\beta}^{[j]})}{1 + \sum_{i=1}^{7}\exp(\mathbf{X}_i^{[j]}\boldsymbol{\beta}^{[j]})}, \quad i = 1, \ldots, 7, \; j = 1, \ldots, 200,$$

is 200 sets of forecasted proportions for the eight genres for the next year (the probability of the baseline category is found by subtracting the sum of the probabilities of the other seven genres from 1). The top panel of Figure 3 shows the boxplots of these forecasted genre proportions for 2019 for the Hot 100 data based on the model chosen in Section 4.1 that did not include time trends. The bottom panel of Figure 3 gives boxplots of the proportions for 2019 for the Pazz & Jop data based on the model selected in Section 4.3.

The Hot 100 boxplots presented convincing and sensible forecasts for genre probabilities for

2019. Rock was predicted to continue its steady decline, and Rap/Hip-Hop continues to rise as a prominent genre, as do Country and Pop. Overall, the forecasted genre probabilities appear reasonable and continue the patterns seen through 2018 which reassures us that this approach to forecasting works well. For the Hot 100, we had year-end genre counts for 2019 which further validated our forecasts. We plotted the observed 2019 proportions with × symbols overlain on the boxplots. As pictured in Figure 3 (top), most of the actual proportions fall near the center of the boxplots. Although the Country proportion is slightly overpredicted and the Pop and Rap/Hip-Hop proportions are slightly underpredicted by the model, the true proportions remain close to the upper and lower quartiles of the distribution of forecasts.

In the Supplementary Material, we provide the analogous boxplots of forecasts for the Hot 100 model with a cubic time trend. This cubic model forecasts the 2019 proportions accurately for many genres; however, it substantially underpredicts the proportions for Pop and Rock. The cubic model incorporates underlying time trends in its forecast, possibly influencing the large forecasted proportion for Rap/Hip-Hop, and the model without time is more heavily influenced by the covariate values and forecasts Country, Pop, and Rap/Hip-Hop to have similar proportions.

The Pazz & Jop boxplots in Figure 3 (bottom) also presented reasonable forecasts for genre probabilities for 2019. The most prominent genres of Rock and Rap/Hip-Hop were predicted to continue their popularity with critics. The forecasts for Country, Folk, and Other were rather uncertain given wide spreads in their boxplots. The model predicts a downturn in popularity among critics for Pop in 2019, potentially due to the slight dip in 2018's observed Pop proportion (see Figure 2), which would manifest itself in the lagged dependent variable in the model to predict the 2019 Pop proportion.

The Village Voice conducted its last Pazz & Jop poll in 2018, but ceased publication after that, so, unlike with the Hot 100 data, we do not have observed Pazz & Jop proportions from 2019 to compare with our forecasted values. *The Village Voice* resumed publication in Spring 2021, so the official Pazz & Jop poll may return in the future. In the meantime, the similarly constructed "Pazz & Jop Rip-off Poll" organized by music insiders Glenn Boothe and Keith Artin (administered since 2019 via a Facebook group) recently attracted media attention (March, 2020; Crawford, 2021, 2022), although we do not include it in our analysis since its voters are volunteers rather than scrupulously selected music critics.

6   Association Between Consumers' and Critics' Preferences

For each year in our data set, consider a 2×8 table whose two rows are the genre counts for that year from the Hot 100 and Pazz & Jop data. Cramér's V (Cramér, 1945), a measure of association of two nominal categorical variables, is defined as $V = [(\chi^2/n)/(\min\{k-1, r-1\})]^{1/2}$ where χ^2 is the usual test statistic for the χ^2 test of association for a matrix with r rows and k columns, with all cell counts totaling n. To measure the association between the general public's and critics' genre preferences from 1974 through 2018, we computed the Cramér's V value between the Hot 100 and Pazz & Jop genre counts for each year and created a time series object with the `ts` function in the `TSA` package (Chan and Ripley, 2020). Note that we used a logit transformation of the Cramér's V data (appropriate since the values fell between 0 and 1). Using the `auto.arima` function in the `forecast` package (Hyndman et al., 2020) to fit an ARIMA model chosen to minimize AIC, we found that an ARIMA(0,1,1) model was the most reasonable fit for the Cramér's V time series. This model choice was supported by the autocorrelation
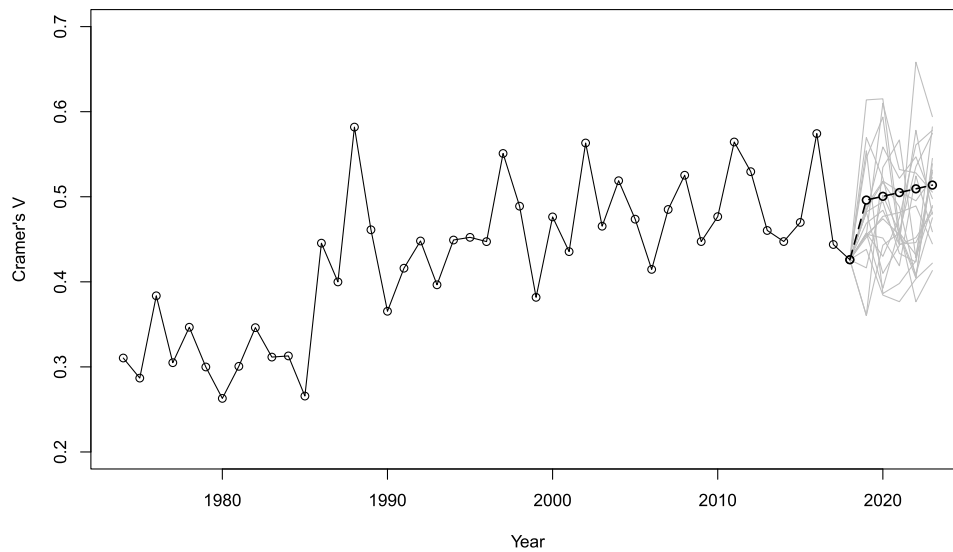
Figure 4: Time series plot of Cramér's *V* values from 1974 through 2018 (solid), and forecasts for 2019 through 2023 (long dash). Simulated individual trajectories are shown in grey.

(ACF) and partial autocorrelation (PACF) functions of the differenced logit-transformed series (see Supplementary Material). The ARIMA(0,1,1) fitted model equation was:

$$W_t = e_t - 0.68e_{t-1},$$

with estimated noise variance 0.073, where W_t is the first differences of the logit-transformed Cramér's *V* series.

Diagnostic plots of the residuals (time series, ACF, and normal Q-Q plots) are provided in the Supplementary Material; these indicated that there was no residual autocorrelation in the model, but the normal errors assumption was not met based on the Q-Q plot and Shapiro-Wilk test (Shapiro and Wilk, 1965) for normality.

From this logit-transformed time series model, we then predicted the association for the next five years (2019 through 2023); the back-transformed Cramér's *V* values are plotted in Figure 4, where 0 represents completely independent genre preferences between consumers and critics and 1 represents completely associated genre preferences. The forecasted values for 2019-2023 are: 0.496, 0.501, 0.505, 0.509, and 0.514. These forecasts predict that on average the association between critics' and consumers' genre preferences over time will slightly increase over the next few years. Over the years, the association between consumer and critical preferences has gradually risen to its current moderate level. This may be due to the demise of Rockism and willingness of critics to be more inclusive about praising commercially popular genres.

Furthermore, with the recent ubiquity of streaming music, critics and consumers may tend to share more similar access to different musical genres than they ever had in the past. The cross-correlation function (CCF) between the "proportion physical" and Cramér's *V* time series provides some support for this speculation. Consider the portions of the two series starting from 1990 (roughly when non-physical formats came into existence). After logit-transforming and prewhitening the two series, there is a modest negative lag-1 correlation of −0.36 (at the boundary of significance), indicating a possible association between the "proportion physical" series and the following year's Cramér's *V* series (see the Supplementary Material for a plot of

the full CCF). Specifically, as physical formats decline and streaming becomes more prevalent, the association between consumers' and critics' preferences tends to rise.

To illustrate the uncertainty in the forecasts, we used our ARIMA(0,1,1) model to simulate 20 future trajectories of Cramér's *V* values, as seen in grey in Figure 4. These were done with the `simulate` function in the `forecast` package (Hyndman et al., 2020) with resampled errors since the normal error assumption was not met. These represent hypothetical realizations of the next five years' Cramér's *V* values, and the forecasts are averages from the population of future realizations.

7   Discussion

Genre, as previously noted, is an imperfect measure for categorizing music due to its inherent subjectivity; however, compared to other measures, genre is one of the most understood and accepted ways to categorize and identify music preference. One of the most difficult aspects of this study was assigning genres to albums where the music crossed stylistic boundaries. (A possible future solution might be found in genre classifiers based on artificial intelligence, extending the methods discussed in Jain et al. (2021).) This difficulty in determining genres was most notable for modern alternative rock that blended pop, rock, and folk influences. In fact, through its development, rock became fragmented into numerous subgenres (Bogdanov et al., 2002). One could conceive of replicating this analysis on rock music alone, treating the subgenres as the categories of interest.

Our analysis presented challenges related to the complexity and autocorrelation of the data, particularly in forecasting. Our forecasting approach yielded honest quantification of the uncertainty through simulating future values of unlagged predictors and through simulating values of our model coefficients. Our model's inclusion of lagged dependent variables (LDVs) as predictors limited our ability to forecast more than one year in the future without having to plug unknown LDV values into the prediction equation. In practice, we would recommend forecasting sequentially one year at a time rather than far into the future.

An alternative approach to fitting two separate multinomial models and then assessing association with Cramér's *V* would be to consider a more complex likelihood that includes the Hot 100 and Pazz & Jop genre counts together and models the correlation as part of the likelihood. However, this approach would preclude using standard models, such as the multinomial logit model, since in such combined data vectors, certain songs could appear in both Hot 100 counts and Pazz & Jop counts at once, violating the independence condition of the multinomial.

To measure genre popularity, we simply counted appearances in the annual Hot 100 and Pazz & Jop lists. An alternative way to measure popularity (and potentially an approach for future research) would be to create a numerical metric that weighted the position in the lists of a song or album. The nature of such a weighting is somewhat arbitrary and would potentially influence heavily the conclusions, so implementing such a scheme would require caution. In addition, since our data are measured annually, the yearly ranking is likely to be influenced by the time of year the song or album was released and thus may not fully reflect its true popularity.

This study demonstrated the changes in music genre popularity over the past 40 years. We modeled how music genres' popularity changes during different decades, noted the importance of new recording technology on genre preference, and measured how critics' opinions related to those of the masses. The decline of Rock and rise of Rap/Hip-Hop, Country, and Pop suggest interesting changes in culture and music preference. Perhaps a shift in generational values is

reflected through changing opinions about lyrics, rhythms, and instrumentation. Understanding what genres are popular and why helps us to relate to others and comprehend the changing cultural environment in which we live, which imbues with new relevance the classic question: What kind of music do you like?

Supplementary Material

This includes the raw data files for both Hot 100 and Pazz & Jop listing singles and albums and genre assignments and all R code used in our analysis along with an explanatory README.txt file. Other items include: Full tables of genre counts for years 1974-2018; Tables of estimated multinomial logit model coefficients; Scatterplot matrix of candidate predictors; Various residual plots; Diagnostic plots for association analysis; Boxplots of 2019 Hot 100 forecasts from the model with the cubic time trend; and the cross-correlation function graph.

Acknowledgement

References

Agresti A (2015). *Foundations of Linear and Generalized Linear Models*. John Wiley & Sons.

Akaike H (1973). Information theory and an extension of the maximum likelihood principle. In: *Second International Symposium on Information Theory* (BN Petrov, F Csáki, eds.), 267–281. Akadémiai Kiadó, Budapest.

Andrews C, Gardiner K, Jain TK, Olomi Y, North AC (2020). Culture, personal values, personality, uses of music, and musical taste. *Psychology of Aesthetics, Creativity, and the Arts.* Advance online publication.

Beckwith J (2016). The evolution of music genre popularity. The DataFace. Online at https://thedataface.com/2016/09/culture/genre-lifecycles. Posted September 7, 2016. Accessed July 8, 2021.

Bogdanov V, Woodstra C, Erlewine ST (2002). *All Music Guide to Rock: the Definitive Guide to Rock, Pop, and Soul.* Hal Leonard Corporation.

Bradlow ET, Fader PS (2001). A Bayesian lifetime model for the "Hot 100" Billboard songs. *Journal of the American Statistical Association*, 96(454): 368–381.

Breslow NE, Clayton DG (1993). Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association*, 88(421): 9–25.

Bryson B (1996). "Anything but heavy metal": Symbolic exclusion and musical dislikes. *American Sociological Review*, 884–899.

Chan KS, Ripley B (2020). *TSA: Time Series Analysis*. R package version 1.3.

Coinnews Media Group LLC, (2021). Historical inflation rates, 1914-2021. https://www.usinflationcalculator.com/inflation/historical-inflation-rates

Condylios S (2020). priceR: Economics and Pricing Tools. R package version 0.1.5.

Cramér H (1945). *Mathematical Methods of Statistics*. Univ. Press, Princeton, NJ.

Crawford S (2021). Facebook's Village Voice Pazz & Jop Rip off Poll 2020, reviewed. Rock NYC. Online at http://rocknyc.live/facebooks-village-voice-pazz-and-jop-rip-off-poll-2020-reviewed.html. Posted January 4, 2021. Accessed February 3, 2022.

Crawford S (2022). Village Voice Pazz and Jop Ripoff Poll 2021. Rock NYC. Online at https://rocknyc.live/village-voice-pazz-and-jop-ripoff-poll-2021.html. Posted January 16, 2022. Accessed February 3, 2022.

Fokianos K, Kedem B (2004). Partial likelihood inference for time series following generalized linear models. *Journal of Time Series Analysis*, 25(2): 173–197.

Greenberg DM, Baron-Cohen S, Stillwell DJ, Kosinski M, Rentfrow PJ (2015). Musical preferences are linked to cognitive styles. *PloS One*, 10(7): e0131151.

GSS Data Explorer (2021). Think of self as liberal or conservative. https://gssdataexplorer.norc.org/variables/178/vshow.

Harvilla R (2017). Have we reached the end of poptimism? *TheRinger. com*, Online at https://www.theringer.com/music/2017/11/16/16666306/taylor-swift-poptimism-2017. Posted November 16, 2017. Accessed June 25, 2021.

Hiller RS, Walter JM (2017). The rise of streaming music and implications for music production. *Review of Network Economics*, 16(4): 351–385.

Hofstede G (2001). *Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations Across Nations*. Sage Publications.

Hosseini R, Le ND, Zidek JV (2012). Time-varying Markov models for binary temperature series in agrorisk management. *Journal of Agricultural, Biological, and Environmental Statistics*, 17(2): 283–305.

Hyden S (2018). *Twilight of the Gods: A Journey to the End of Classic Rock*. HarperCollins.

Hyndman R, Athanasopoulos G, Bergmeir C, Caceres G, Chhay L, O'Hara-Wild M, et al. (2020). forecast: Forecasting functions for time series and linear models. R package version 8.13.

Jain R, Sharma R, Nagrath P, Jain R (2021). Music genre classification chatbot. In: *Proceedings of Second International Conference on Computing, Communications, and Cyber-Security*, 393–408. Springer.

Liang KY, Zeger SL (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, 73(1): 13–22.

Liu M, Hu X, Schedl M (2018). The relation of culture, socio-economics, and friendship to music preferences: A large-scale, cross-country study. *PloS One*, 13(12): e0208186.

Macrotrends (2021a). U.S. GDP growth rate 1961-2021. https://www.macrotrends.net/countries/USA/united-states/gdp-growth-rate.

Macrotrends (2021b). U.S. national unemployment rate. https://www.macrotrends.net/1316/us-national-unemployment-rate.

March K (2020). Village Voice Pazz & Jop Rip-off Poll's best albums of 2019. Flood Magazine. Online at https://floodmagazine.com/73002/village-voice-pazz-jop-rip-off-polls-best-albums-of-2019/. Posted January 8, 2020. Accessed February 3, 2022.

North AC, Hargreaves DJ (2007). Lifestyle correlates of musical preference: 1. relationships, living arrangements, beliefs, and crime. *Psychology of Music*, 35(1): 58–87.

Pedrero-Esteban LM, Barrios-Rubio A, Medina-Ávila V (2019). Adolescentes, smartphones y consumo de audio digital en la era de spotify. *Comunicar*, 27: 103–112.

Powers D (2013). *Writing the Record: The Village Voice and the Birth of Rock Criticism*. University of Massachusetts Press.

R Core Team (2020). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria.

Recording Industry Association of America (2021a). U.S. recorded music revenues by format. https://www.riaa.com/u-s-sales-database/.

Recording Industry Association of America (2021b). U.S. sales database. https://www.riaa.com/u-s-sales-database/.

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.

Richards C (2015). Do you want poptimism? or do you want the truth? The Washington Post. Online at https://www.washingtonpost.com/entertainment/music/at-the-top-of-the-pop-music-heap-theres-no-criticizing-the-view/2015/04/16/d98d53a8-e1f2-11e4-b510-962fcfabc310__story.html?utm_term=.66c297e6a6f4. Posted April 17, 2015. Accessed June 25, 2021.

Robins W (2020). Missing the Pazz & Jop critics poll. *PsAudio Copper*, 102: 32–35. Online at https://www.psaudio.com/wp-content/plugins/psa-copper-magazine/public/pdf/issue/issue-102.pdf. Accessed June 25, 2021.

Sanneh K (2004). The rap against rockism. *The New York Times*, 31, 351–354. Published October 31, 2004.

Schwarz G (1978). Estimating the dimension of a model. *Annals of Statistics*, 6(2): 461–464.

Shapiro SS, Wilk MB (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3/4): 591–611.

Silver D, Lee M, Childress CC (2016). Genre complexes in popular music. *PloS One*, 11(5): e0155471.

Spivack S, Philibotte SJ, Spilka NH, Passman IJ, Wallisch P (2019). Who remembers the Beatles? the collective memory for popular music. *PloS One*, 14(2): e0210066.

Vlegels J, Lievens J (2017). Music classification, genres, and taste patterns: A ground-up network analysis on the clustering of artist preferences. *Poetics*, 60: 76–89.

Warton DI, Hui FK (2011). The arcsine is asinine: the analysis of proportions in ecology. *Ecology*, 92(1): 3–10.

Woolhouse M, Bansal J (2013). Work, rest and (press) play: Music consumption as an indicator of human economic development. *Journal of Interdisciplinary Music Studies*, 7: 45–71.

Yee TW (2021). VGAM: Vector Generalized Linear and Additive Models. R package version 1.1-5.