PANDEMICS, MUSIC, AND COLLECTIVE SENTIMENT: EVIDENCE FROM THE OUTBREAK OF COVID-19

Meijun Liu[1] **Eva Zangerle**[2] **Xiao Hu**[1]
Alessandro Melchiorre[3,4] **Markus Schedl**[3,4]
[1] The University of Hong Kong, Hong Kong SAR, China
[2] University of Innsbruck, Innsbruck, Austria
[3] Johannes Kepler University Linz, Austria
[4] Linz Institute of Technology (LIT), Linz, Austria
Correspondence:xiaoxhu@hku.hk

ABSTRACT

The COVID-19 pandemic causes a massive global health crisis and produces substantial economic and social distress, which in turn may cause stress and anxiety among people. Real-world events play a key role in shaping collective sentiment in a society. As people listen to music daily everywhere in the world, the sentiment of music being listened to can reflect the mood of the listeners and serve as a measure of collective sentiment. However, the exact relationship between real-world events and the sentiment of music being listened to is not clear. Driven by this research gap, we use the unexpected outbreak of COVID-19 as a natural experiment to explore how users' sentiment of music being listened to evolves before and during the outbreak of the pandemic. We employ causal inference approaches on an extended version of the LFM-1b dataset of listening events shared on Last.fm, to examine the impact of the pandemic on the sentiment of music listened to by users in different countries. We find that, after the first COVID-19 case in a country was confirmed, the sentiment of artists users listened to becomes more negative. This negative effect is pronounced for males while females' music emotion is less influenced by the outbreak of the COVID-19 pandemic. We further find a negative association between the number of new weekly COVID-19 cases and users' music sentiment. Our results provide empirical evidence that public sentiment can be monitored based on collective music listening behaviors, which can contribute to research in related disciplines.

1. INTRODUCTION

Music listening has various functions in people's daily lives, especially in terms of psychological aspects. Mood management and regulation are two major psychological uses of music listening [31, 41–43] and have been widely investigated in previous studies [18, 21, 40, 44]. In the literature, there are mainly two kinds of music mood being addressed. One is "expressed" mood, the other is "induced" mood. The former refers to the mood that is intended to be expressed by a piece of music whereas the latter refers to listeners' emotional state or feeling induced by listening to a music piece or situations with which people associate a song [5, 58]. Both kinds of music mood have been extensively studies in the field of Music Information Retrieval (MIR) in recent decades [59]. It is well known that people often search and select music based on the mood of music that fits their current (emotional) needs [12] and it has been found that people's mood is responsive to events in daily life [10, 25]. However, the influence of real-world events, such as human or natural disasters, on users' mood-based music selection remains less explored, especially in the scale of collective choices of users in a society.

As of May 4, 2020, with 3.57 million cases and over 250 thousand deaths reported, COVID-19 has posed a severe threat to public health. To contain the virus, multiple measures have been taken, such as city lock-downs, social distancing, travel restrictions, and university closures, disrupting people's daily life and routine. High unemployment and the economic damages caused by COVID-19 make people suffer from increasing economic stress [17, 24]. Causing economic and social pressure, the pandemic is putting enormous stress on all of us and might trigger feelings of distress and anxiety [8, 56]. Large-scale disasters are often accompanied by increases in depression, a broad range of other psychological stress and behavioral disorders [16, 20, 34]. Investigating whether and the extent to which collective sentiment inferred from music listening behavior is influenced by the pandemic can not only deepen our understanding of the relationship between real-world events and users' sentiment reflected by the music they listen to, but also contribute to mitigating negative mental health impacts caused by the pandemic, and help people adapt, and be resilient during distress times.

Inspired by the research gap and driven by the emerging pandemic, this study aims to explore the association between real-world events and the sentiment of music being listened to based on users' music listening history at

© M. Liu, E. Zangerle, X. Hu, A. Melchiorre,and M. Schedl . Licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0). **Attribution:** M. Liu, E. Zangerle, X. Hu, A. Melchiorre,and M. Schedl . "Pandemics, music, and collective sentiment: evidence from the outbreak of COVID-19", 21st International Society for Music Information Retrieval Conference, Montréal, Canada, 2020.

a large scale. In this study, we use the term *users' music sentiment* to refer to the aggregated collective sentiment of music listened to by a population of users. We use the difference in differences (hereafter DD) method [2], a widely used causal inference approach, and base our analysis on an extension of the LFM-1b dataset [45]. This dataset provides us with a substantial set of listening histories of users from around the world as well as user-generated tags that we derive sentiment information from. We use the unanticipated outbreak of COVID-19 as a natural experiment to investigate the causality between pandemics and collective sentiment reflected by music people choose to listen to.

Specifically, we aim to answer the following research questions:

RQ1: How did the first COVID-19 case in a country affect the music sentiment of users in that country?

RQ2: How did the number of new COVID-19 cases and hence, the spread of the disease in a country affect the music sentiment of users in that country?

Using the DD model, our analyses can reveal causal relationship between real-world events, a pandemic in our case, and users' music sentiment. Results of this study could shed light on designs of music recommenders to be more sensitive to real-life events. More importantly, our results provide convincing evidence that collective sentiment is hampered by the pandemic. Beyond the field of MIR, findings of this study could provide empirical evidence on the extent to which public sentiment could be monitored by music listening behaviors of a concerned population. This could contribute to science in related disciplines such as social psychology, sociology, and journalism.

2. RELATED WORK

Two streams of research show a possible relationship between real-world events and users' sentiment of music being listened to, while the causal pathway remains unexplored. Links between real-worlds events and people's mood have been widely documented in previous research [50, 51]. However, to our best knowledge, there has been no study on the association between real-world events and users' collective sentiment as reflected by music they listen to.

2.1 Real-world events linked to mood

Mood indicates a set of transient, fluctuating affective states in terms of individuals' feelings [4, 26]. A variety of factors are related to individuals' mood state, including personality [14], ongoing events, experiences, or the environmental milieu [57]. Real-world events or the daily experiences, conceptualized as situational or contextual factors, have long been recognized as important determinants of daily mood in abundant research [50, 51].

Pioneering studies on the association between daily events and mood are often based on self-reported data on the same day. For example, Lewinsohn and colleagues [29] found a negative correlation between the number of pleasant events and depression by exploiting subjects' self-reported data on daily experiences and self-ratings of mood states. Similar results have been found by Rehm [36] and Stone [49, 51]. In particular, Stone found a same-day association between major daily events and mood, based on self-reported events and moods of 50 men [51]. However, studies based on self-reported data in the same-day context are not sufficient to demonstrate a causal relationship because mood could impact the reliability of event reporting [9]. More convincing evidence of the causal relationship between real-world events and mood could be provided based on longitudinally self-reported data on events and moods over a period of time.

The occurrences of events have an impact on changes in positive and negative mood, respectively. Pleasant events are normally related to positive mood, such as exercise [13, 28], family, friends, and leisure time [50] and other pleasant daily events [27]. There is also evidence of the links between various negative daily events and negative mood, such as interpersonal conflict [6], negative interpersonal interactions [38], stressful work-related events [38], daily hassle [11], undesirable daily events [1], and other daily stressors [15, 52]. For instance, using the data extracted from diaries, Zuckerman demonstrated that young adults reported significantly higher levels of depression when they reported interpersonal conflicts [7].

Current research on the effect of major real-world events on mood is limited [47]. The majority of previous work is built on the self-reported data of small samples. On one hand, it is unclear whether small samples were sufficiently representative; on the other hand, the self-rating of mood state can be subjective.

2.2 Real-world events and music listening

Schedl et al. [46] took an initial step to explore the correlation between real-world events and music consumption behavior by leveraging listening events from Last.fm and world-wide events from Google Trends. Performing an intervention time series analysis, the authors found that changes in listening behavior might be correlated to real-world events. However, only two variables were taken into account: the number of events identified by Google Trends and the absolute number of listening events. Whether or not real-world events are linked to users' music sentiment remains uncovered.

Inspired by this research gap , this study aims to examine the association between real-world events (the COVID-19 pandemic) and the collective sentiment reflected in people's music listening behaviors. Based on (an extension of) the LFM-1b dataset, one of the largest datasets of music listening histories available to date, this study examines the relationship between the outbreak of COVID-19 and the collective music sentiment.

3. DATA AND METHODOLOGY

In the following, we present the data utilized and subsequently, we describe the methods underlying our analyses.

3.1 Data

We built upon the LFM-1b dataset [45] to explore the relationship between the outbreak of the COVID-19 pandemic and users' music sentiment. More precisely, we gathered the listening records of users in the LFM-1b dataset between November 1, 2019, one month before the first COVID-19 case in the world was confirmed in China [1], and March 27, 2020. Even though the outbreaks are still ongoing in many places as of the writing of this paper, this time frame captures the times of first confirmed COVID-19 cases in all countries involved in this study. We obtained more than 28 million listening events shared during this period of time, 21 weeks in total, generated by 12,278 Last.fm users from 40 countries from the LFM-1b dataset. From the Our World in Data website, [2] we collected data on confirmed COVID-19 new cases by date for these countries. In particular, we identified the date of the first confirmed COVID-19 case in each country and define it as the date when the COVID-19 outbreak started in this country.

3.2 Measuring music sentiment

In Last.fm, tracks and artists are associated with tags created by users. To capture the sentiment of music listened to by our users, we collect the artists whose music pieces are included in our dataset, that is, have been listened to by included users during the concerned period of time. Although tracks also have tags, track-level tags indicating sentiment might be too sparse as there are a much larger number of tracks than artists. Therefore, using artist sentiment (as opposed to track sentiment) provides us with a substantially larger set of tags with sentiment. The sentiment of an artist can then be calculated by methods described in the next paragraphs.

We capture the sentiment values assigned to all artists by crawling the user-created tags assigned to those artists from Last.fm. Subsequently, we follow Zangerle et al.'s [60] approach for the computation of sentiment values based on the tags. Specifically, we utilize sentiment lexica, a widely used unsupervised sentiment detection method for the extraction of sentiment information from the tags. A sentiment lexicon is a list of words, where each word is assigned a sentiment value (e.g., on a scale from 0 to 1, describing a range of sentiments from negative to positive) [54]. We rely on a set of widely used dictionaries that have been shown to provide the best coverage and accuracy according to Ribeiro's benchmark of sentiment dictionaries [39]: AFINN [32], Opinion Lexicon [22], SentiStrength [55], and Vader [23]. The different lexica provide different notions of polarity and strength of polarity. Hence, we normalize the set of sentiments to a range between 0 and 1 by using linear min-max feature scaling.

The following steps are taken to perform the sentiment computation based on these dictionaries. First, we employ whitespace as delimiters to tokenize the tags as tags may consist of several words such as in "nothing left to fear". We then use the tokens as the input to the matching step between tokens and dictionaries. To do so, we apply lemmatization to the tokens contained in the sentiment dictionaries (utilizing the lemmatization method of the Python's NLTK package). Next, we match each tag token to the set of dictionaries. For each matched token, we extract the assigned sentiment value. In the case that a token matches multiple entries of the dictionaries (i.e., if it is contained in multiple dictionaries), we utilize the arithmetic mean of those values as the sentiment value of the token. After having computed the sentiment value for each token, we assign the tag the mean of the sentiment values of all tokens contained in the original tag.

The sentiment of an artist is then defined as the weighted average sentiment values of the tags assigned to the artist, as shown in Equation 1:

$$AS_j = \sum_{i=1}^{n} v_i \cdot \frac{f_i}{F} \tag{1}$$

where (AS_j) stands for the sentiment value of the j^{th} artist; v_i denotes the sentiment value of the i^{th} tag for artist j; f_i refers to the frequency of tag i for this artist, and F denotes the total frequency of tags for the artist.

We consider this a reasonable estimate as the tags applied to each artist reflect the collective sentiment a large number of listeners had towards the artist over a long period of time. It can thus be assumed that tags are not biased by certain (subgroups of) users or any short term events.

3.3 Difference in differences approach

To capture the relationship between the outbreak of the COVID-19 pandemic and the sentiment of music listened to by users, the DD model is used [2]. This method measures the differential effect of certain changes on the dependent variable of treatment groups versus that of control groups [19]. One distinct advantage of the DD model is that it can disclose causality. In this study, the concerned change is the outbreak of COVID-19. The dependent variable is the users' music sentiment extracted from their listening behaviors (cf. Section 3.2). In this study, for a given period of observation (i.e., a specific week), users in countries with confirmed COVID-19 cases are in the treatment group, whereas users in countries without confirmed cases form the control group.

Besides, to control for users' other characteristics that may affect their sentiment of music, we introduce fixed effects regression, the major method for regression analysis of panel data. It is an extension of multiple regression that exploits panel data to control for variables that differ across entities but are constant over time [48]. This allows controlling unobserved changes during our observation period. We incorporate user gender, age, and country as control variables.

The DD model employed for the COVID-19 pandemic is shown in Equation 2.

$$\begin{aligned} Y_{it} = &\alpha + \beta_1 \left(I_{it}\right) + \beta_2 Gender_i + \\ &\beta_3 Age_i + C_i + D_t + \epsilon_{it} \end{aligned} \tag{2}$$

where Y is a dependent variable, I are the independent variables, it indicates observations for user i in week t and ϵ_{it} indicates the error term in the equation. We detail the utilized dependent and independent variables in

[1] There are rumors on suspected cases earlier than November 1, 2019 at different places around the world. In this study, we only consider cases officially reported in mainstream media.

[2] `https://ourworldindata.org/covid-cases`

Sections 3.4 and 3.5, respectively. Furthermore, the equation also contains controls and fixed-effect variables as described above. Specifically, fixed-country effects (C_i in the equation) allow controlling the time-invariant, country-specific characteristics of users as e.g., socio-cultural background may be related to users' artist listening behavior. The observable changes of global events that might influence users' listening behavior worldwide can be controlled by adding a week fixed effect (D_t). Controls of gender and age of users are also included in Equation 2.

The coefficient for the dependent variable is an effective difference-in-differences estimate of the average impact of the COVID-19 outbreak on the sentiment of music listened to by users. The effect of the outbreak of COVID-19 on user's music sentiment in a week is estimated using the Ordinary Least Squares method, which is appropriate for the continuous dependent variable.

3.4 Dependent variables

We use two dependent variables to build two models, by which we can cross check robustness of the results. Both are on user's music sentiment calculated based on the measure of music sentiment described in Section 3.2. The first dependent variable is the average sentiment of artists a user listened to (hereafter referred to as*USE*) in a given week. The values of USE range from 0 to 1. The larger the USE value is, the more positive the user's music sentiment. USE captures the average sentiment of artists the user listened to in each day, weighted by the relative consumption frequency of the artists, as shown by Equation 3:

$$USE = \sum_{i=1}^{n} AS_i \cdot \frac{c_i}{C} \tag{3}$$

where AS_i denotes the sentiment value of the i^{th} artist the user listened to on the given day; c_i indicates the playcount (number of times the user listened to artist i); and C denotes the total playcount of the user on that day. Based on daily USE values, we then aggregate this user's daily USE to the weekly level through averaging. The weekly average USE of a user is the first dependent variable.

The second dependent variable we propose to use is whether the music sentiment of a user in a given week is extremely positive (i.e., whether or not the weekly USE value of a user is in the 90^{th} percentile of USE values of all users in our dataset). We refer to this variable as *POS*.

3.5 Independent variables

For our DD analyses, we propose to use the outbreak of COVID-19, denoted as *COVID-19* hereafter, as the independent variable. Specifically, it is a binary variable indicating whether or not the first COVID-19 case has been confirmed in a given country in a given week. The variable is 0 for observations (a.k.a samples) before the outbreak in a country and 1 for observations/samples after the outbreak. In addition, to explore the association between the number of new COVID-19 cases and user's music sentiment, we also include the weekly number of new COVID-19 cases in a country as the second independent variable, denoted as *COVID19 cases*.

Variable	Mean	Std. Dev.	Min	Max
USE	0.3332	0.0782	0	0.9813
POS	0.0976	0.2967	0	1
COVID-19	0.2038	0.4028	0	1
COVID-19 cases	0.8344	1.9249	0	10.0542
Age	32.3674	13.1637	6	126
Gender	0.1731	0.3783	0	1

Table 1. Descriptive statistics of variables modeled in this study; N =184,277 (weekly observations of 12,278 users for 21 weeks); gender indicates male (0) or female (1).

Table 1 summarizes the descriptive statistics of dependent and independent variables in this study. The distribution of USE for the samples is presented in Figure 1.

3.6 Interaction effect between the COVID-19 and gender

To explore whether there is a gender difference regarding the effect of the COVID-19 pandemic on user' music sentiment, we generate two interaction terms between gender and the independent variables respectively, *COVID#Gender* and *Case#Gender*. Specifically, we multiply COVID-19 and COVID-19 cases with gender respectively and add each of the interaction terms to Equation 2 for separate modeling.

Besides, after running regression models on all sampled users, we conduct DD analysis for females and males separately, to further examine possible gender differences.

In sum, for each dependent variable (USE, POS), we run four regression models on each independent variable (COVID-19 and COVID19 cases): the original one indicated by Equation 2, the one with added interaction term, and two on female and male users only.

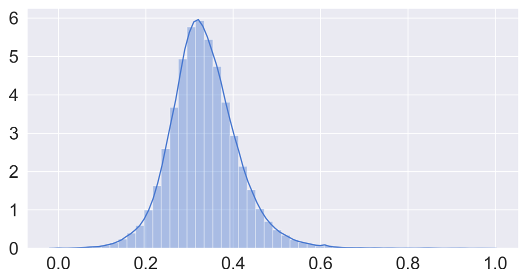

Figure 1. Distribution of USE (average sentiment of artists the user listened to), with USE on x-axis and probability density on y-axis.

4. RESULTS

4.1 COVID-19 outbreak and users' music sentiment

The effects of the independent variable, COVID-19, on dependent variable, weekly USE (user's music sentiment) are shown in columns 1 to 4 in Table 2 (Panel 1). We observe a significant and negative association between the outbreak of COVID-19 in a country and users' music sentiment (USE): The coefficient on COVID-19 for USE is -0.002 ($p < 0.01$, t-test) in column 1 of Table 2 (Panel 1), suggesting that the sentiment values of the users in countries with COVID-19 outbreak is 0.002 lower than that of the users in countries where COVID-19 has not appeared yet. In other words, compared to sentiment of artists listened to by users in countries without COVID-19, that of users exposed to this pandemic is more negative.

In the USE model, the coefficient for Gender in column 1 (Panel 1) is 0.014 at the significance level of 0.01, indicating that the sentiment of artists listened to by females is more positive than that of males. The coefficient for age is significantly negative in column 1 (Panel 1), suggesting that younger users' average weekly music sentiment is more positive than older users, though the difference is small and nearly close to zero.

Gender differences are also found in the model of USE with the intersection term (column 2, Panel 1). With a value of 0.0025 ($p < 0.05$), the coefficient on the interaction term between COVID-19 outbreak and gender, COVID#Gender, is significantly positive, indicating that females are less influenced by the pandemic, as compared to their male peers. In column 2 of Table 2 (Panel 1). The estimated USE for females and males before and after the outbreak of COVID-19 is shown in Figure 2. The figure suggests a very slight decrease in USE after the outbreak of COVID-19 for females, as compared to a clear decline in users' USE for males. The regression results are consistent based on the models that only include either female (column 3, Panel 1) or male users (column 4, Panel 1): there is no significant impact of COVID-19 outbreak on USE in the model of females (column 3), while there is a significantly negative effect for males as shown in column 4 of Table 2 (Panel 1).

The results of the analyses of the POS variable (i.e., whether music sentiment of users is extremely positive) provide consistent findings that user's music sentiment turns more negative after the outbreak of COVID-19. Column 5 of Table 2 (Panel 1) reveals that there is a significantly negative relationship between the outbreak of COVID-19 and the positiveness of users' music sentiment (POS). After the outbreak of COVID-19, the probability that users' weekly averaged music sentiment reaches extremely positive values decreased by nearly 2.8% ($p < 0.01$). The coefficient on the interaction term between COVID-19 and gender is not significant (column 6, Panel 1). Columns 7 and 8 in Panel 1 both show that the likelihood that the weekly USE for females and males get extremely positive declines by 1.6% ($p < 0.05$) and 3.2% ($p < 0.01$), respectively.

4.2 COVID-19 cases and users' music sentiment

Users' sentiment of music is negatively associated with the number of COVID-19 cases in a country. In other words, in a country where a larger number of people are infected with COVID-19 in a given week, the weekly average sentiment of artists the user listened to becomes more negative. From column 1 in Table 2 (Panel 2), one unit growth in COVID-19 new cases in the week leads to a decrease of 0.05 in weekly USE of users , i.e., 64.10% (i.e., 0.05 divided by standard deviation of USE) standard deviation of users' USE. Again, we find the negative effect of COVID-19 cases only for males, as shown in column 4 of Table 2 (Panel 2). For the dependent variable POS, results in column 5 of Table 2 (Panel 2) shows that, one unit increase in COVID-19 new cases in the week is related to a 67% lower probability that the weekly user sentiment reaches extremely positive values. There is no gender difference regarding the effect of new COVID-19 cases on users' POS as the interaction term between new COVID-19 cases and Gender is not significant (column 6, Panel 2), and the coefficients of COVID-19 cases on POS for females and males are very close, as shown in columns 7 and 8 of Table 2 (Panel 2).

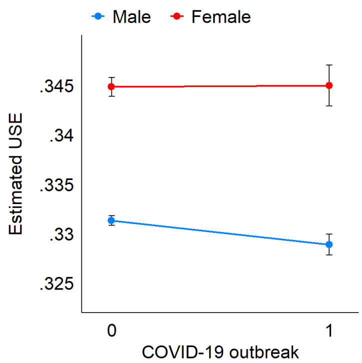

Figure 2. Estimated interaction effect between COVID-19 outbreak and gender.

5. DISCUSSION

Our results show that, after the outbreak of COVID-19 in a country, the weekly sentiment of artists users listened to turns more negative. These results are robust regardless of whether the dependent variable is the weekly average sentiment value of users' artist listening (USE) or whether or not users' music sentiment is extremely positive (POS). The findings hence show significant causality between the real-world event (i.e., the COVID-19 pandemic) and users' music sentiment. We find that, after the first COVID-19 case in a country was confirmed, the weekly sentiment of artists users listened to decreased by 0.002, nearly 3% of standard deviation of users' USE (column 1 Panel 1 in Table 2. This negative effect is pronounced for males, whereas females' music emotion is less influenced by the outbreak of the COVID-19 pandemic. We further find a negative association between the number of COVID-19 new cases in the current week and users' music sentiment: one unit growth in COVID-19 new cases in a given week brings a decrease of 0.05, i.e., 64.10% of standard deviation of user's weekly music sentiment (column 1, Panel 2 in Table 2. One unit increase in COVID-19 new cases in a week is related to a 67% lower probability that user's weekly music sentiment reaches extremely positive (column 5, Panel 2 in Table 2).

We provide convincing evidence to show a negative impact of public health crisis on collective sentiment using the unexpected outbreak of COVID-19 pandemic as

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel 1				
Model	**USE**				**POS**			
	All	All	Female	Male	All	All	Female	Male
COVID19	-0.0020*** (0.0006)	-0.0024*** (0.0006)	0.0008 (0.0015)	-0.0028*** (0.0007)	-0.0284*** (0.0023)	-0.0290*** (0.0024)	-0.0161*** (0.0062)	-0.0316*** (0.0025)
Gender	0.0141*** (0.0005)	0.0136*** (0.0005)			0.0335*** (0.0018)	0.0329*** (0.0020)		
COVID#Gender		0.0025** (0.0012)				0.0036 (0.0047)		
Age	-0.0000** (0.0000)	-0.0000** (0.0000)	0.0001** (0.0000)	-0.0000*** (0.0000)	-0.0001 (0.0001)	-0.0001 (0.0001)	0.0003* (0.0002)	-0.0001** (0.0001)
Constant	0.3321*** (0.0005)	0.3322*** (0.0005)	0.3428*** (0.0013)	0.3328*** (0.0005)	0.0999*** (0.0020)	0.1000*** (0.0020)	0.1199*** (0.0053)	0.1025*** (0.0020)
Observations	184,277	184,277	31,894	152,383	184,277	184,277	31,894	"152,383"
R-squared	0.030	0.030	0.036	0.025	0.014	0.014	0.020	0.011
				Panel 2				
Model	**USE**				**POS**			
	All	All	Female	Male	All	All	Female	Male
COVID19 cases	-0.0005** (0.0002)	-0.0005** (0.0002)	-0.0003 (0.0006)	-0.0006** (0.0002)	-0.0067*** (0.0008)	-0.0065*** (0.0009)	-0.0069*** (0.0023)	-0.0068*** (0.0009)
Gender	0.0141*** (0.0005)	0.0141*** (0.0005)			0.0340*** (0.0018)	0.0351*** (0.0020)		
Case#Gender		0.0001 (0.0003)				-0.0014 (0.0010)		
Age	-0.0000** (0.0000)	-0.0000** (0.0000)	0.0001** (0.0000)	-0.0000*** (0.0000)	-0.0001 (0.0001)	-0.0001 (0.0001)	0.0003* (0.0002)	-0.0001** (0.0001)
Constant	0.3321*** (0.0005)	0.3322*** (0.0005)	0.3432*** (0.0013)	0.3327*** (0.0006)	0.0996*** (0.0020)	0.0995*** (0.0020)	0.1224*** (0.0055)	0.1017*** (0.0021)
Observations	184,277	184,277	31,894	152,383	184,277	184,277	31,894	152,383
R-squared	0.030	0.030	0.036	0.025	0.013	0.013	0.020	0.011

Table 2. Estimated effect of the outbreak of COVID-19 (Panel 1) and number of new COVID-19 cases (Panel 2) on users' music sentiment. Robust standard errors are in parentheses; gender indicates female (1) or male (0). All models employ country and week-fixed effects (*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$). For columns (1) to (4), the dependent variable is USE. From columns (5) to (8), the dependent variable is POS. Columns (1) and (5) report the regression results based on the models that include control variables, i.e., gender and age for all sampled users; columns (2) and (6) indicates the regression results based on the models with control variables and an interaction term between COVID-19/Case and gender; columns (3), (4) and columns (7), (8) show the results based on the models with either female or male users.

an example that ensures a causal relationship through the different of differences method. Wide spread outbreaks of infectious diseases often cause psychological distress and symptoms of mental illness [3, 35]. Our results suggest that the average sentiment of artists users listened to becomes more negative after the first COVID-19 case was confirmed in a country, which is consistent with reported detrimental effect of pandemics on public mental health [30, 33, 37]. We further find that in countries where more new COVID-19 are confirmed, users' music sentiment turns more negative. During this acute public health crisis, people might experience fears of infection and worry about the pandemic's consequences. The disruption of daily routines and the social isolation imposed by the "stay at home orders" adopted in many countries also cause compounding personal stress and anxiety. The unemployment and financial losses caused by physical isolation may also strengthen feelings of distress. These negative emotions, as shown by our results, have been reflected by the sentiment of artists users listened to during the pandemic.We find that females' music sentiment is less influenced by the outbreak of COVID-19, which is probably due to the gender differences in socio-economic roles that females might be less financially stressed than males [53].

6. CONCLUSION

This study applies the difference of differences method to find that, after the first COVID-19 case in a country was confirmed, the weekly sentiment of artists users listened to became more negative. This negative effect is pronounced for males. We further find a negative association between the number of new COVID-19 cases in a week and users' music sentiment. The results could provide useful implications for the design of music recommendation systems during public health crises or other disasters, which could help manage users' mood, enhance mental health, and mitigate negative psychological impacts of pandemics. The findings also provide empirical evidence that large-scale aggregation of music listening data can help monitor collective sentiment of listener populations.

7. ACKNOWLEDGEMENT

This study is supported by National Natural Science Foundation of China (No. 61703357), and a General Research Fund from the Research Grants Council of the Hong Kong S. A. R., China (No. HKU 17607018).

8. REFERENCES

[1] Glenn Affleck, Howard Tennen, Susan Urrows, and Pamela Higgins. Person and contextual features of daily stress reactivity: individual differences in relations of undesirable daily events with mood disturbance and chronic pain intensity. *Journal of Personality and Social Psychology*, 66(2):329, 1994.

[2] Joshua D Angrist and Jörn-Steffen Pischke. *Mostly harmless econometrics: An empiricist's companion*. Princeton university press, 2008.

[3] Yanping Bao, Yankun Sun, Shiqiu Meng, Jie Shi, and Lin Lu. 2019-ncov epidemic: address mental health care to empower society. *The Lancet*, 395(10224):e37–e38, 2020.

[4] Bonnie G Berger and Robert W Motl. Exercise and mood: A selective review and synthesis of research employing the profile of mood states. *Journal of applied sport psychology*, 12(1):69–92, 2000.

[5] Kerstin Bischoff, Claudiu S Firan, Raluca Paiu, Wolfgang Nejdl, Cyril Laurier, and Mohamed Sordo. Music mood and theme classification-a hybrid approach. In *ISMIR*, pages 657–662, 2009.

[6] Niall Bolger, Anita DeLongis, Ronald C Kessler, and Elizabeth A Schilling. Effects of daily stress on negative mood. *Journal of personality and social psychology*, 57(5):808, 1989.

[7] Niall Bolger and Adam Zuckerman. A framework for studying personality in the stress process. *Journal of personality and social psychology*, 69(5):890, 1995.

[8] Wenjun Cao, Ziwei Fang, Guoqiang Hou, Mei Han, Xinrong Xu, Jiaxin Dong, and Jianzhong Zheng. The psychological impact of the covid-19 epidemic on college students in china. *Psychiatry research*, page 112934, 2020.

[9] David M Clark and John D Teasdale. Diurnal variation in clinical depression and accessibility of memories of positive and negative experiences. *Journal of abnormal psychology*, 91(2):87, 1982.

[10] Lee Anna Clark and David Watson. Mood and the mundane: Relations between daily life events and self-reported mood. *Journal of personality and social psychology*, 54(2):296, 1988.

[11] Anita DeLongis, Susan Folkman, and Richard S Lazarus. The impact of daily stress on health and mood: psychological and social resources as mediators. *Journal of personality and social psychology*, 54(3):486, 1988.

[12] J Stephen Downie and Sally Jo Cunningham. Toward a theory of music information retrieval queries: System design implications. 2002.

[13] Kathryn M Fritz and Patrick J O'Connor. Acute exercise improves mood and motivation in young men with adhd symptoms. *Medicine and science in sports and exercise*, 48(6):1153–1160, 2016.

[14] Scott K Fry and Bernd G Heubeck. The effects of personality and situational variables on mood states during outward bound wilderness courses: An exploration. *Personality and individual differences*, 24(5):649–659, 1998.

[15] Shelly L Gable, Harry T Reis, and Andrew J Elliot. Behavioral activation and inhibition in everyday life. *Journal of personality and social psychology*, 78(6):1135, 2000.

[16] Sandro Galea, Arijit Nandi, and David Vlahov. The epidemiology of post-traumatic stress disorder after disasters. *Epidemiologic reviews*, 27(1):78–91, 2005.

[17] Anuj Gangopadhyaya and A Bowen Garrett. Unemployment, health insurance, and the covid-19 recession. *Health Insurance, and the COVID-19 Recession (April 1, 2020)*, 2020.

[18] Walter Gantz, Howard M Gartenberg, Martin L Pearson, and Seth O Schiller. Gratifications and expectations associated with pop music among adolescents. *Popular Music & Society*, 6(1):81–89, 1978.

[19] Paul J Gertler, Sebastian Martinez, Patrick Premand, Laura B Rawlings, and Christel MJ Vermeersch. *Impact evaluation in practice*. The World Bank, 2016.

[20] Emily Goldmann and Sandro Galea. Mental health consequences of disasters. *Annual review of public health*, 35:169–183, 2014.

[21] David J Hargreaves and Adrian C North. The functions of music in everyday life: Redefining the social in music psychology. *Psychology of music*, 27(1):71–83, 1999.

[22] Minqing Hu and Bing Liu. Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 168–177. ACM, 2004.

[23] Clayton J Hutto and Eric Gilbert. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Eighth international AAAI conference on weblogs and social media*, 2014.

[24] Wolfram Kawohl and Carlos Nordt. Covid-19, unemployment, and suicide. *The Lancet Psychiatry*, 7(5):389–390, 2020.

[25] Peter Kuppens, Zita Oravecz, and Francis Tuerlinckx. Feelings change: Accounting for individual differences in the temporal dynamics of affect. *Journal of personality and social psychology*, 99(6):1042, 2010.

[26] Andrew M Lane and Peter C Terry. The nature of mood: Development of a conceptual model with a focus on depression. *Journal of applied sport psychology*, 12(1):16–33, 2000.

[27] Christopher A Langston. Capitalizing on and coping with daily-life events: Expressive responses to positive events. *Journal of Personality and Social Psychology*, 67(6):1112, 1994.

[28] Arthur R LaPerriere, Michael H Antoni, Neil Schneiderman, Gail Ironson, Nancy Klimas, Panagiota Caralis, and Mary Ann Fletcher. Exercise intervention attenuates emotional distress and natural killer cell decrements following notification of positive serologic status for hiv-1. *Biofeedback and Self-regulation*, 15(3):229–242, 1990.

[29] Peter M Lewinsohn and Julian Libet. Pleasant events, activity schedules, and depressions. *Journal of abnormal psychology*, 79(3):291, 1972.

[30] Nianqi Liu, Fan Zhang, Cun Wei, Yanpu Jia, Zhilei Shang, Luna Sun, Lili Wu, Zhuoer Sun, Yaoguang Zhou, Yan Wang, et al. Prevalence and predictors of ptss during covid-19 outbreak in china hardest-hit areas: Gender differences matter. *Psychiatry research*, page 112921, 2020.

[31] Adam J Lonsdale and Adrian C North. Why do we listen to music? a uses and gratifications analysis. *British Journal of Psychology*, 102(1):108–134, 2011.

[32] Finn Årup Nielsen. A new anew: Evaluation of a word list for sentiment analysis in microblogs. *arXiv preprint arXiv:1103.2903*, 2011.

[33] Paul C Perrin, O Lee McCabe, George S Everly, and Jonathan M Links. Preparing for an influenza pandemic: mental health considerations. *Prehospital and disaster medicine*, 24(3):223–230, 2009.

[34] Robert H Pietrzak, Melissa Tracy, Sandro Galea, Dean G Kilpatrick, Kenneth J Ruggiero, Jessica L Hamblen, Steven M Southwick, and Fran H Norris. Resilience in the face of disaster: prevalence and longitudinal course of mental disorders following hurricane ike. *PLoS One*, 7(6), 2012.

[35] Jianyin Qiu, Bin Shen, Min Zhao, Zhen Wang, Bin Xie, and Yifeng Xu. A nationwide survey of psychological distress among chinese people in the covid-19 epidemic: implications and policy recommendations. *General psychiatry*, 33(2), 2020.

[36] Lynn P Rehm. Mood, pleasant events, and unpleasant events: Two pilot studies. *Journal of Consulting and Clinical Psychology*, 46(5):854, 1978.

[37] Dori B Reissman, Patricia J Watson, Richard W Klomp, Terri L Tanielian, and Stephen D Prior. Pandemic influenza preparedness: adaptive responses to an evolving challenge. *Journal of Homeland Security and Emergency Management*, 3(2), 2006.

[38] Rena L Repetti. Short-term effects of occupational stressors on daily mood and health complaints. *Health Psychology*, 12(2):125, 1993.

[39] Filipe N Ribeiro, Matheus Araújo, Pollyanna Gonçalves, Marcos André Gonçalves, and Fabrício Benevenuto. Sentibench-a benchmark comparison of state-of-the-practice sentiment analysis methods. *EPJ Data Science*, 5(1):1–29, 2016.

[40] Keith Roe. Swedish youth and music: Listening patterns and motivations. *Communication Research*, 12(3):353–362, 1985.

[41] Suvi Saarikallio and Jaakko Erkkilä. The role of music in adolescents' mood regulation. *Psychology of music*, 35(1):88–109, 2007.

[42] Thomas Schäfer and Peter Sedlmeier. From the functions of music to music preference. *Psychology of Music*, 37(3):279–300, 2009.

[43] Thomas Schäfer and Peter Sedlmeier. What makes us like music? determinants of music preference. *Psychology of Aesthetics, Creativity, and the Arts*, 4(4):223, 2010.

[44] Thomas Schäfer, Peter Sedlmeier, Christine Städtler, and David Huron. The psychological functions of music listening. *Frontiers in psychology*, 4:511, 2013.

[45] Markus Schedl. The LFM-1B Dataset for Music Retrieval and Recommendation. In *Proceedings of the 2016 ACM on International Conference on Multimedia Retrieval*, ICMR '16, pages 103–110, New York, NY, USA, 2016. ACM.

[46] Markus Schedl, Eelco Wiechert, and Christine Bauer. The effects of real-world events on music listening behavior: An intervention time series analysis. In *Companion Proceedings of the The Web Conference 2018*, WWW '18, pages 75–76, Republic and Canton of Geneva, Switzerland, 2018. International World Wide Web Conferences Steering Committee.

[47] Alexander J Shackman, Jennifer S Weinstein, Stanton N Hudja, Conor D Bloomer, Matthew G Barstead, Andrew S Fox, and Edward P Lemay Jr. Dispositional negativity in the wild: Social environment governs momentary emotional experience. 2017.

[48] James H Stock and Mark W Watson. *Introduction to econometrics*. Prentice Hall, 3rd edition, 2015.

[49] Arthur A Stone. The association between perceptions of daily experiences and self-and spouse-rated mood. *Journal of Research in Personality*, 15(4):510–522, 1981.

[50] Arthur A Stone. Event content in a daily survey is differentially associated with concurrent mood. *Journal of Personality and Social Psychology*, 52(1):56, 1987.

[51] Arthur A Stone and John M Neale. Effects of severe daily events on mood. *Journal of personality and social psychology*, 46(1):137, 1984.

[52] Arthur A Stone, John M Neale, and Saul Shiftman. Daily assessments of stress and coping and their association with mood. *Annals of Behavioral Medicine*, 15(1):8–16, 1993.

[53] Joanna Syrda. Spousal relative income and male psychological distress. *Personality and Social Psychology Bulletin*, page 0146167219883611, 2019.

[54] Maite Taboada, Julian Brooke, Milan Tofiloski, Kimberly Voll, and Manfred Stede. Lexicon-based methods for sentiment analysis. *Computational linguistics*, 37(2):267–307, 2011.

[55] Mike Thelwall, Kevan Buckley, Georgios Paltoglou, Di Cai, and Arvid Kappas. Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology*, 61(12):2544–2558, 2010.

[56] Cuiyan Wang, Riyu Pan, Xiaoyang Wan, Yilin Tan, Linkang Xu, Cyrus S Ho, and Roger C Ho. Immediate psychological responses and associated factors during the initial stage of the 2019 coronavirus disease (covid-19) epidemic among the general population in china. *International journal of environmental research and public health*, 17(5):1729, 2020.

[57] David Watson. *Mood and temperament*. Guilford Press, 2000.

[58] Zhongzhe Xiao, Emmanuel Dellandrea, Weibei Dou, and Liming Chen. What is the best segment duration for music mood analysis? In *2008 International Workshop on Content-Based Multimedia Indexing*, pages 17–24. IEEE, 2008.

[59] Yi-Hsuan Yang and Homer H Chen. Machine recognition of music emotion: A review. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 3(3):40, 2012.

[60] Eva Zangerle, Chih-Ming Chen, Ming-Feng Tsai, and Yi-Hsuan Yang. Leveraging affective hashtags for ranking music recommendations. *IEEE Transactions on Affective Computing*, 2018.