Online Appendices for *Disperse and Preserve the Perverse: Computing How Hip-Hop Censorship Changed Popular Music Production in China*

APPENDIX A. Fieldwork

My fieldwork also proves that Hip-Hop musicians were heavily struck by the censorship in general, regardless of their popularity. I did interviews with 28 Hip-Hop musicians, producers, and managers in the summer of 2019 in 5 Chinese cities (Beijing, Shanghai, Chengdu, Changsha, Wuhan) where there is a prominent local Hip-Hop scene. Some of them enjoyed national fame, while others were relatively unknown to the mainstream audience. As I was told, none of them received any official notifications regarding the censorship; they all heard about it from the media coverage of the press conference. But they suffered from it in very concrete ways: many of them lost their branding contracts, the opportunities to show up in mainstream television programs, and the permission to hold concerts at large public venues due to the censorship for at least 6 months. Although they were not banned from publishing songs on online music platforms, they started to pay more attention to their use of language, avoiding explicitly problematic expressions. Some of them turned to make a more pop-oriented, melodic style of Hip-Hop music that they wished this "safer" style of music can avoid censorship and reach a broader audience. Most of them believed that the censorship would eventually go away, and they seemed right as the second season of *The Rap of China*, this time under a different Chinese name (which was changed literally from "China Has Hip-Hop" to "China's New Rap" while the official English name remained the same), was aired in July 2018 in midst of public doubts about its broadcasting. Some Hip-Hop musicians were even featured in CCTV programs again later in that year. The release of the censorship seems to signal that Hip-Hop music has been changed to the extent that the cultural regulators don't need to be concerned about it anymore.

Appendix B. Data Collection

B.1. Main dataset

There are several reasons to collect a dataset as such. First, the dataset incorporates Chinese Language Songs, which include songs made by artists not only from mainland China but also from Hong Kong, Taiwan, Malaysia, Singapore, and other Chinese-speaking regions. While many of these non-mainland-Chinese musicians are willing to profit from mainland China's large market, they have to work with publishers or online music platforms located in mainland China to have their songs legally circulated there (Baranovitch 2003). Consequently, their songs are also subject to cultural censorship in mainland China, and they are supposed to be impacted by censorship as well. Moreover, while the lyrics of most songs are written in standard Mandarin Chinese, which is the official language in mainland China, some are written in other Chinese Languages such as Cantonese. They are nevertheless also subject to cultural censorship.

Second, I collected songs from studio albums instead of singles or compilation albums primarily for the balance between accessibility and generality. The configuration of the Chinesemusic.com website restricts the number of items allowed to be displayed. The number of singles released in each year between 2015 and 2018 is higher than allowed to be displayed so that I am not able to collect a significant proportion of the singles. On the other hand, there are too few compilation albums and they are usually released only by relatively prominent artists, which would make the dataset extremely biased and limited. Since songs released as singles or in the compilation albums are usually included in the studio albums and the number of albums does not exceed that allowed to be displayed, songs from official studio albums are best to represent all the songs released in each year.

Third, besides Hip-Hop songs, I also collected songs of three other genres: Pop, Rock, and Folk. The genres are labeled by musicians themselves, so they indicate how musicians categorize their own music. When uploading albums to the platform, musicians are asked to choose one major genre from 24 genre tags available on the platform[1] for the album. Each genre also has its respective subgenre tags (for example, there are 24 subgenre tags under Hip-Hop, such as Pop Rap and Trap Rap). Musicians can also choose other genre tags and subgenre tags as the album's minor genre, yet the album will be classified into the category of the major genre. Once the album genre is decided, all the songs in the album will be tagged in the same way.

The three genres I choose are representing different types of genre compared to Hip-Hop. Pop, in fact, is technically not a genre in the same sense as Hip-Hop (Lena & Peterson 2008).

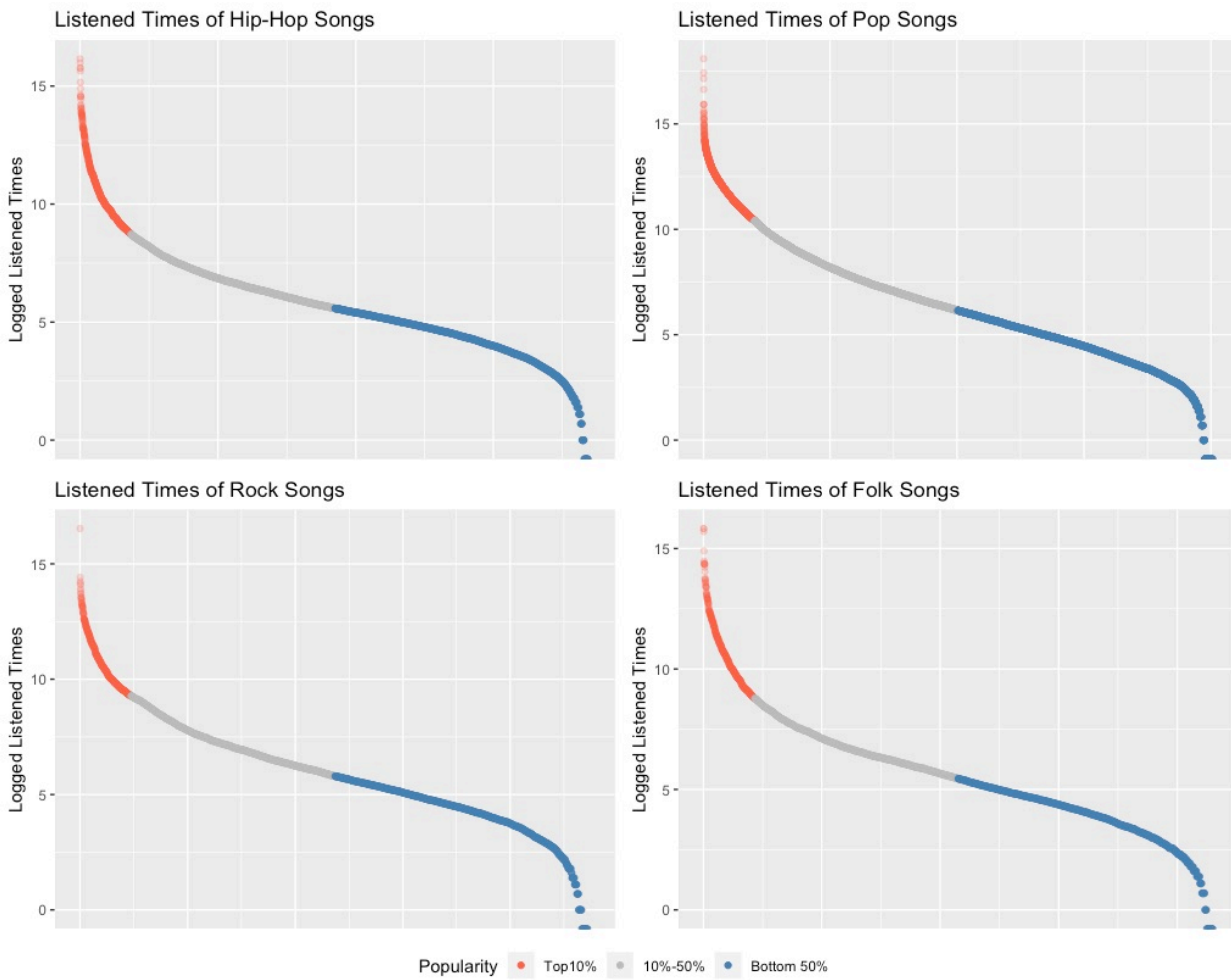
[1] The 24 genre tags are Pop (流行), Rock (摇滚), Folk (民谣), Electronic (电子), R&B (节奏布鲁斯), Jazz (爵士), Light Music (轻音乐), Hip-Hop (嘻哈[说唱]), ACG (动漫), Blues (布鲁斯), Metal (金属), Punk (朋克), World Music (世界音乐), New Age (新世纪), Country (乡村), Raggae (雷鬼), Classical (古典), Singer/Songwriter (唱作人), Latin (拉丁), Chinese Characteristic (中国特色), Experimental (实验), Children (儿童), Audio Book (有声书), Stage & Screen & Entertainment (舞台/银幕/娱乐).

The distinction between Pop and other genres lies primarily in the business-operational aspect as the former is targeting a broader and less niched audience. Rock and Folk are similar to Hip-Hop as they also target niche markets respectively, yet Rock is musically more affinitive to Hip-Hop than Folk. Historically, Hip-Hop musicians are known to collaborate more with Rock musicians than Folk musicians (Baranovitch 2003). Also, there are more music genres that combine Hip-Hop with Rock than Folk in my dataset: there are 143 songs with genre tags that include Hip-Hop and Rock and 35 of them are without Pop tag (including Pop Rap and Pop Rock), while 91 songs have genre tags that include Hip-Hop and Folk and only 1of them are without Pop tag (including Pop Rap and Folk Pop). This demonstrates Hip-Hop's affinity to Rock relative to Folk. In general, this shows that Pop is a relatively low-contrast category (Hannan 2010; Hannan et al. 2007) that is flexible with mixing with other genres, and Hip-Hop, Rock, and Folk are all relatively high-contrast categories that are niched. The contrast level of genre and its affinity to each other have further implications on how they react to cultural censorship, which will be elaborated on in the analysis part.

Fourth, the main focus of the analysis will be all songs released in 2017 and 2018, which cover the four periods based on existing reports and my fieldwork, each of which lasts around half a year: Period I starts from Jan 1 to Jun 24, 2017, when the first season of The Rap of China was aired; it is followed by Period II, which ends on Jan 19, 2018, when the censorship was announced; Period III went on until Jul 14, 2018, when the second season of The Rap of China was broadcasted; Period IV covers the rest of the year 2018. I also collected songs released in 2015 and 2016 for training a Hip-Hop classifier, which will be further used to measure the probability of a 2017/2018 song being classified as Hip-Hop if it were released in 2015/2016. This allows comparisons between songs across time in terms of their sound quality. I argue that 2015/2016 songs will serve as an ideal training set for genre classifier as they will neither confuse the classifier by absorbing the possible change that I indeed want to measure in 2017 and 2018 (which will happen if using 2017/2018 songs instead) nor will they severely affected by the endogenous innovations and changes developed within each genre (which will happen if using much earlier songs instead).

GRAPHIC B.1 shows the distribution of the times each 2017/2018 song was listened to at the time of data collection, descending from most listened to least listened from left to right. The value of the y-axis is the natural log of the actual value due to the skewness of the distribution for the sake of clarity. The red dots are the top 10% listened-to songs of each genre, while the blue ones are the bottom 50%. The graphic generally supports that, although the choice of taking the top 10% listened-to songs and the bottom 50% listened songs as "popular songs" and "unpopular songs" is arbitrary, they make sense intuitively as the top 10% songs generally have been listened more than ca. 20,000 times (e^{10}=22026) and the bottom 50% songs less than 150 times (e^{5}=148).

GRAPHIC B.1. Listened Times of Songs of Four Genres



B.2. Sample of censored songs

The songs were located using an internet archive service, Wayback Machine (web.archive.org). Wayback Machine is a digital library that provides free public access to collections of digitized materials, including historical snapshots of websites. There are multiple snapshots of Chinesemusic.com in 2017 and 2018 collected in Wayback Machine, from which a limited amount of information on the platform at the time of snapshot is accessible. Although the snapshots did not capture all the songs released before the censorship, it did capture user-generated playlists that were most "liked" by other users, to which some censored songs were affiliated. Therefore, I identified 11 most-liked Hip-Hop playlists and compared their snapshots before and after the censorship in the Wayback Machine to locate 116 songs that were pulled down after the censorship. For each censored song, I used various sources to collect and verify its basic information, and I eventually managed to collect the lyrics and the audio files of 104 songs. if possible. [2] These songs are used to train a censorship classifier which measures how similar a given song is compared to the censored songs acoustically.

B.3. Other limitations of the dataset

The dataset also has some other limitations that are worth noting but, I contend, will not significantly change the results of the analysis. Due to competition for music copyrights between online music platforms, Chinesemusic.com might not be able to provide access to songs whose copyright is owned exclusively by other music platforms. Some artists, regardless of the matter of copyright, might also upload their songs exclusively to other platforms due to their own preferences. I argue that while this may affect the generality of the dataset in representing the Chinese music market, it will not do significant harm to our investigation of the censorship effect because such effect is supposed to be similar across platforms since the censorship is targeting at the entire music market rather than a specific music platform. Moreover, the dataset does not include songs that are produced and circulated privately, so it is not representing every possible song produced in and for mainland China. That being said, these songs have little to do with censorship since they are not publicly circulated and thus are of little interest to the study here. Last but not least, while the date of a song being uploaded to the platform is commonly available, we don't know exactly when the musician started and finished the song. The production cycle of an album varies greatly among musicians - from days to years – and there might also be a lag between the production and release of a song. Yet this will not affect our conclusion regarding music production since we would expect songs initially finished before the censorship would be re-considered by the musicians if they were to release the songs after the

[2] The sources I used include youtube.com. rapzh.com, muxiv.net, and flac123.com.

censorship: they would either maintain its original form, revise it to try to conform to the censorship or delete it from the album, all of which shows an effort of re-production of the song. Also, technological development in music production is allowing musicians to make songs and release them in days. This is especially the case for Hip-Hop music, where song production can be completed by a single person with devices as simple as a personal laptop and a microphone, which is also confirmed by the musicians that I interviewed in my fieldwork.

Appendix C. Methods

C.1. Audio

The approach used to construct acoustic measurements for this paper is inspired by Askin & Mauskapf's (2017) study of over 25,000 Billboard songs, which provides a new way of studying the musical features of a massive amount of songs. The authors used the service from a music intelligence company, The Echo Nest, to extract acoustic features provided by the service, such as "acousticness", "danceability", and "energy". These features are generated based on Music Information Retrieval (MIR) techniques. Although The Echo Nest would have been a useful instrument for this study, it has been unavailable to the public since its acquisition by Spotify in 2014. Fortunately, there are multiple MIR libraries that are open source, which indicates a sustainable way of doing similar research. The one used here is LibROSA, a python package for music and audio analysis, which similarly extracts acoustic features of the music and is widely used in the MIR community (McFee et al. 2015).

MIR is particularly useful for music genre classifications by comparing patterns of acoustic features of a song in question with those of songs whose genre is pre-coded. It predicts a song's genre based on how similar its pattern is compared to the typical pattern of songs of each genre, which can be extracted from studying massive pre-coded songs through machine learning algorithms. While the task for this study is not to predict a song's genre accurately, similar logic can be used to "predict" how similar a given song is compared to a group of pre-coded songs. Specifically, the task here is to see how similar or dissimilar the songs produced after the censorship are compared to those produced before or pulled down by the censorship.

To make such a comparison, I construct two classifiers using machine learning algorithms.[3] The first one is a music genre classifier, calculating the probability of a song to be classified into each genre. The acoustic features that I extracted using LibROSA are commonly used in MIR-based music genre classification, including chroma frequencies, spectral centroid, spectral roll-off, zero-crossing rate, and Mel-frequency cepstral coefficients (MFCC) (Hamel and Eck 2010; Trohidis et al. 2008; Xu, Maddage, and Shao 2005). I use all songs released in 2015 and 2016 as the training set for the classifier, recoding the four genres into a binary Hip-Hop versus Non-Hip-Hop variable for the interest of this study and the performance of the classifier.[4] The classifier will calculate the probability of a given song being classified into either Hip-Hop or Non-Hip-Hop based on the pattern of the acoustic features. Since the sum of the two probabilities will always be 1, I will hence focus solely on the probability of Hip-Hop, a score between 0-1indicating how "Hip-Hopy" the predicted song is, or how possible the predicted song is a Hip-Hop song if it were released in 2015 or 2016. I name the score *Hip-Hopiness*, which will be the key dependent variable for the study. I use a similar method to construct the second classifier - censorship classifier - which is trained based on the censored songs.[5] The classifier will predict how similar a given song sounds to censored songs. I name the probability score *Censorshipness*, the other key dependent variable. I used the two classifiers to score the *Hip-Hopiness* and *Censorshipness* for each song released in 2017 and 2018. GRAPHIC C.1 shows the distribution of the two dependent variables among all songs in 2017 and 2018:

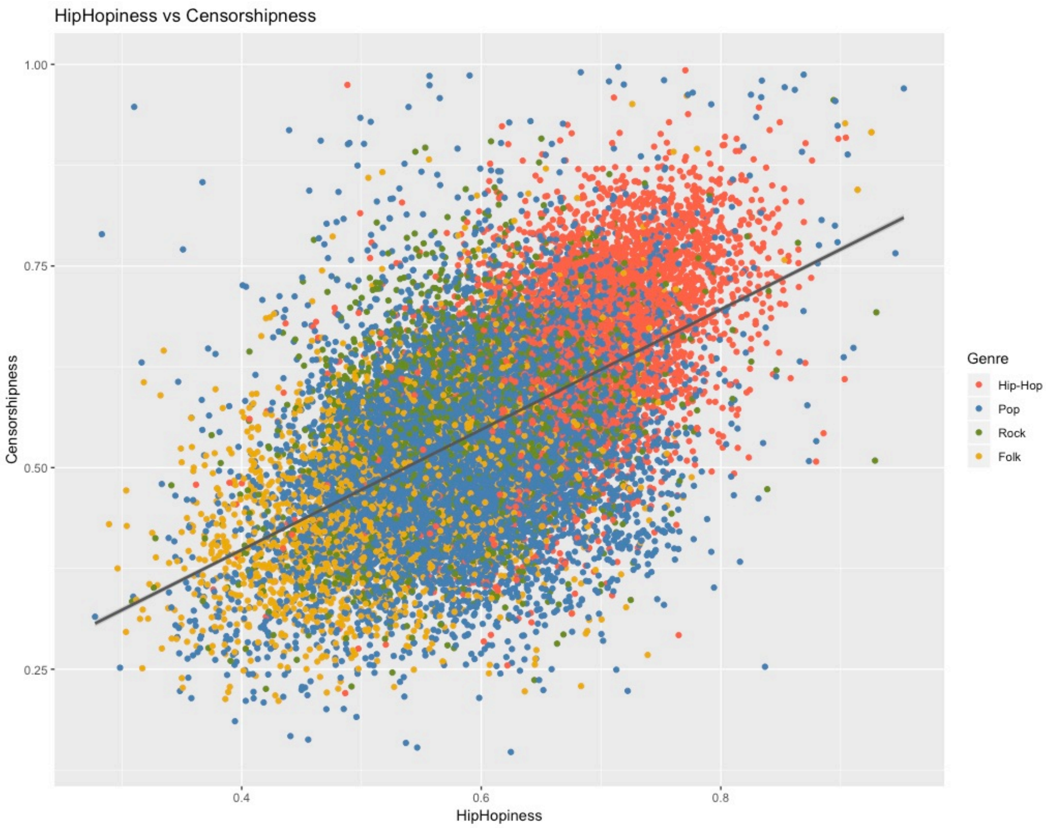
GRAPHIC C.1. Hip-Hopiness and Censorshipness of 2017 & 2018 Songs

---

[3] I use neural networks for this study among all machine learning algorithms because neural networks are useful for recognizing the pattern of the input, which will be particularly suitable here.

[4] Classifiers can be trained based on different metrics, such as accuracy. When using accuracy metrics, the classifier will be trained to maximize the number of times it predicts the right class over the total number of predictions. It is particularly useful when there are a smaller number of classes and the sample is balanced across classes. Since there are 2262 songs labeled as *Hip-Hop* in 2015 and 2016, I recoded all the *Pop*, *Rock*, and *Folk* songs in 2015 and 2016 as *Non-Hip-Hop* and randomly sampled 2200 songs to balance the training set. The training set thus consists of 4462 songs pre-coded as *Hip-Hop* and *Non-Hip-Hop*. I split the whole training set in 4:1 ratio where the classifier will be trained based on 80% of the sample (training sample) and tested in the rest 20% of the sample (testing sample). Eventually, the classifier scores a prediction accuracy of 77.26% for the training sample and 75.52% for the testing sample.

[5] Since there are 104 audio files of the censored songs available, I randomly sampled 120 songs from all the uncensored songs to construct a balanced training set. I similarly used neural networks and the same extracted features for training the classifier. Eventually, the classifier scores an 80.85% accuracy in predicting the censored song correctly.



Each of the colored dots in GRAPHIC C.1 represents one song released in 2017 or 2018 with its color representing the major genre labeled on the music platform, and the black line is the linear regression of *Censorshipness* against *Hip-Hopiness*. The graphic shows that Hip-Hop songs in 2017 and 2018 are generally having high *Hip-Hopiness* and high *Censorshipness* scores, and Folk songs are having mostly low scores for both, while many Pop and Rock songs are sandwiched in the middle. This meets our expectations and indicates that the two classifiers are well-performing. There is also a significantly positive correlation between *Hip-Hopiness* and *Censorshipness*, which makes sense since all the censored songs in the training set are Hip-Hop songs. However, the dots are relatively dispersed around the regression line. This indicates that there are many songs with high *Hip-Hopiness* but relatively low *Censorshipiness*, implying that the censored songs have their distinctive traits in terms of their acoustic characteristics compared to other Hip-Hop songs.

GRAPHIC C.2. Hip-Hopiness of Hip-Hop songs, 2017-2018

GRAPHIC C.2 shows in an intuitive way what the models in the main article imply. In this example of the *Hip-Hopiness* of 17/18 *Hip-Hop* songs, each colored dot represents one song with its value on the Y-axis referring to its score of *Hip-Hopiness* and the X-axis referring to the song's release date. The colors denote in which of the four periods were the songs released, while the three vertical dashed lines represent the three interventions. The four horizontal black lines are the linear regressions on *Hip-Hopiness* by release date in each period. The dots are omitted in all the graphics in the main article for the clarity of illustration.

C.2. Structural Topic Models

Like other topic models, Structural Topic Model (STM) identifies topics in a large corpus by constructing a generative model of word counts but, unlike others, also incorporates document-level information into its model, recognizing that document-level information may also affect the way document is written (Roberts, Stewart, and Tingley 2014). The topics identified by the STM model are groups of words that are associated under a single theme discovered in the processes of modeling. Human interpreters have to make sense of the association between the words to make sense of that topic. In other words, topics are not automatically labeled but dependent on human interpretation. This in fact is an advantage for this study as the model acknowledges the variety of meanings of terms across different contexts and leaves the meaning to human interpreters, which is similar to how audiences make sense of the lyrics (DiMaggio, Nag, and Blei 2013).

The key dependent variable in our examination of lyrics is the topic *prevalence*, i.e., how much of a document is associated with a topic. With STM, we can explore how topic prevalence is a function of document metadata. In this case, I am interested in how the prevalence of topics identified by the STM model that can be interpreted as sensitive to regulators is changed by the censorship. Specifically, I estimate the parameter of the covariate which denotes whether the song is released before or after the censorship. Similar to my analysis of the music, I focus primarily on the shift of topic prevalence from Period II to Period III brought by the censorship within each popularity group (Top 10% and Bottom 50%) of each genre (Hip-Hop, Pop, Rock, Folk). The model allows me to interpret how the censorship changes the prevalence of a topic of interest in, say, Top 10% Hip-Hop songs released before and after the censorship.

In this study, I use the stm R package to run structural topic modeling (Roberts et al. 2014). The stm R package provides handy ways to configure the modeling. Before modeling, researchers need to specify multiple parameters, including the number of topics that the model will generate (k), the type of the model initialization, and the model for topic prevalence which includes covariates that researchers believe to be relevant. The package designers recommend using spectral initialization, which is deterministic and globally consistent under reasonable conditions (Roberts et al. 2014). Also, they suggest that 60-100 topics work well with a corpus of 10k to 100k documents. Taking their recommendations, I use spectral initialization in my model and set the k to be 100. As for the model for topic prevalence, I include variables that indicate whether the song is released before or after the censorship, which genre is it labeled as, and whether it belongs to the Top 10% or Bottom 50% popularity group. Among all 100 topics generated by the model, I identify 5 topics that are of particular interest to this study from the first 6 words that associated with the topic of highest probability, shown in TABLE C.1:

TABLE C.1. Five stm Generated Topics

Topic	Most probable words	Typical song lyrics
Sex	got, go, beauti, money, psycho, pussi	Ca$h Flow, *Money*
Violence	起来 [become], 暴力 [violence], 焦虑 [stressed], 四个 [four], 世界 [world], 烧 [burn]	五石散 & 1999, *噩梦侦探*
Smoke & Drink	兄弟 [brother], 一起 [together], 挺 [quite], 抽烟 [smoke], 哥们 [brother], 喝 [drink]	满舒克, T-T, & Toy 王奕, *Me & Ma Bros*
Politics	美丽 [beautiful], 草原 [grassland], 中国 [china], 最美 [the most beautiful], 祖国 [motherland], 家乡 [hometown]	锅包肉 & QC-琴橙, *巍巍华夏*
Struggle	life, fight, die, live, way, know	Greatfly & 嫩桃弟弟, *YoungRichChigga*

I also tried different *k*s between 30-200 and found the model works well with *k* between 80-120, generating topics of my interest consistently with largely identical words within those topics. I also asked two research assistants, both of whom are native Mandarin speakers, to validate that the labels make sense to human interpreters. This confirms the validity of our findings in the main article.

C.3. Dictionary

Musicians usually re-write sensitive words that are explicitly related to sensitive topics to avoid being directly detected and pulled down by the platform. When musicians upload their song lyrics to the platform, the lyrics are usually reviewed by platform algorithms and human regulators before they can be published on the site. The length of the review period varies across songs; yet if the lyrics are considered sensitive, they will not be allowed to publish, and the musicians will be notified and asked to upload them again after revision. The words that are usually re-written include those that express sexual behavior (e.g., fuck), indicate drug-use and violence (e.g., marijuana), or refer to political figures (e.g. 主席 [Chairman]). It is worth noting that musicians tend to re-write words both in Chinese and English since sensitive terms in both languages will be targeted by platform algorithms. Sometimes, Chinese musicians also write their lyrics in languages other than Chinese or English (e.g., French), yet the number of these songs is small, and the musicians don't necessarily write in those languages to engage with sensitive topics or to use sensitive terms in that language.

The pattern of when and how to re-write sensitive terms is, to say the least, messy. Since the criteria of what the platform will censor are usually black-boxed, musicians have to make decisions based on whether they believe the term will be censored. Therefore, there are cases where the terms are written in their original form in some song lyrics but are re-written in others. Also, the way musicians re-write sensitive terms varies greatly. For example, they may re-write the term with symbols or special characters (e.g., "f**k" for "fuck"), using the acronym of the *pinyin* (the official romanization system for Standard Chinese in mainland China) of the Chinese word (e.g., "sb" for "傻屄" [*sha bi*, a common Chinese swear word meaning literally "stupid cunt"]), or changing the term to another similar-sounding term (which may also change to a different language, e.g., "法克" [a literally meaningless word sounding *fa ke*] for "fuck") by musicians or by the platform with the consent of the musicians. Sometimes there are multiple versions to re-write the same term (e.g., f**k, f*ck, fvck, fk for "fuck"). These covering strategies are widely used in the lyrics, which allow them still to be distributed publicly online. In some cases, however, lyrics with uncovered sensitive terms are still able to be distributed on the platform, indicating the randomness of the cultural regulators. This brings great difficulties in identifying what the sensitive terms are and how they are covered.

To deal with the task, I first looked for all the terms that contain symbols (e.g., asterisks) and left out those that are not used for re-writing (e.g., for splitting sections). For those terms without any indications of what they are (e.g., "***"), I searched the internet to see if there is an uncensored version of the lyrics and if not, I listened to the song to figure out what the terms are. I then looked for terms that do not contain symbols but are spelled in an alternative way. I identified these terms by randomly sample a batch of 100 songs from the dataset for 50 times and see if there are terms that are re-written as something else, and I focused primarily on the terms identified above. I also compare lyrics of censored songs with those of uncensored to validate my findings: if a term is re-written in uncensored songs but not in the censored songs, then the term is very likely to be sensitive. After checking and searching the terms back and forth, I identified the following sensitive terms in TABLE C.2, annotated with different ways in which they are re-written:

TABLE C.2. The Dictionary of Sensitive Terms

Sensitive Term	Ways of re-written (examples)
fuck/fucking/motherfucker	f**k, f*k, fxxk, fxxx, funk, fvk, fux, motherxxxxx, mutherfucka, fkcu, fk, 法克, 马泽法克, 吗的法克
shit	sh*t, s***, sxxt, sh!t, 谢
bitch	bi*ch, b***h, b*tch, bit*h, bc, b7, 碧池
pussy	pus*y, p***y, pxssy, p****
ass/asshole	as*hole
suck	s*ck, suxk, s**k
dick	d*ck, d
rape	reap
make love	
damn	d*mn, d**n
weed	w*ed, wed, w***
marijuana	xxxx
dope	d**e
nigger/chigger	ni**a, chi**a
government	g0vernment
鸡巴/玑玈［dick］	鸡*, **, g8
屄 [pussy]	逼, b
屌 [dick, dope]	吊, diao, d
龟头 [glans]	**
肏 [fuck]	操, 草, 艹, cao, 日
大麻 [weed]	**

贩毒 [sell drugs]	贩*
毒品/毒药 [drugs]	**, dp, d 品
瘾君子 [drug addict]	**子
吸毒 [take drugs]	**
嗑药 [take drugs]	**
鸦片 [opium]	**
罂粟 [poppy]	yingsu, **
杀人/枪杀/杀戮 [murder]	**
春药 [philter]	**
婊子 [bitch]	*子, 表子
打炮 [have sex]	**
打飞机 [masturbate]	打**
炮王 [the king of having sex]	**
做爱 [make love]	**
前戏 [foreplay]	**
破处 [lose virginity]	破*
强奸 [rape]	
阴道 [vagina]	**
骚货 [tart]	*货
艳舞 [sex dance]	**
色情 [obscene]	
意淫 [sexual fantasy]	意 y
胸器 [boobs]	**
底裤/内裤 [underwear]	xx
大便/粪便/屎 [shit]	大*, **
屁股 [ass]	**

脑残 [retarded]	**
弱智 [retarded]	**
他妈的/他娘的 [fucking]	tmd, ***, 卡玛的
滚你妈 [fuck off]	***
干你娘的 [motherfucking]	*****的
去你妈的 [motherfucking]	****
放屁 [bullshitting]	**
政府 [government]	正府
共产党 [the Communist Party]	***, **党
主席 [chairman]	**
条子 [cop]	**
喇嘛 [lama]	**
人权 [human rights]	**
动乱 [unrest]	**
言论控制 [speech control]	****

Appendix D Regression Tables

The tables below show the coefficients and standard errors (in the parentheses) of the interrupted time series regressions described in the main article. To mitigate the concern about heteroskedasticity in time series data, I also tested all the models using heteroskedasticity-consistent standard errors, known as "HC3" developed by MacKinnon & White (1985) and recommended to use when heteroskedasticity is under concern (Long and Ervin 2000), and check the value and significance of the coefficients. The results hold for all the regressions and are hence not reported here.

TABLE D.1. Interrupted Times Series Regression on Hip-Hopiness Across Genre

Interrupted Time Series Regression on Hip-Hopiness

	Dependent variable:			
	Hip-Hopiness (%)			
	Hip-Hop	*Pop*	*Rock*	*Folk*
Date	0.004	-0.007***	0.002	-0.016***
	(0.004)	(0.002)	(0.004)	(0.005)
I-ROC	0.125	-0.140	-0.630	0.172
	(0.511)	(0.307)	(0.481)	(0.681)
day-Post-ROC	-0.006	0.011***	-0.004	0.020***
	(0.005)	(0.003)	(0.005)	(0.006)
I-Censor	-1.283***	-1.163***	2.864***	-0.013
	(0.455)	(0.273)	(0.548)	(0.656)
day-Post-Censor	0.016***	0.005*	-0.012**	0.001
	(0.004)	(0.003)	(0.005)	(0.006)
I-ROC2	-0.795*	-0.347	1.443**	0.950
	(0.424)	(0.280)	(0.565)	(0.710)
day-Post-ROC2	-0.020***	-0.002	0.013**	-0.023***
	(0.004)	(0.003)	(0.006)	(0.007)
Constant	66.771***	59.676***	58.746***	56.763***
	(0.483)	(0.257)	(0.441)	(0.544)
Observations	6,374	16,702	4,346	4,001
R^2	0.004	0.007	0.012	0.008
Adjusted R^2	0.003	0.007	0.010	0.007
Residual Std. Error	6.748 (df = 6366)	6.406 (df = 16694)	6.057 (df = 4338)	7.728 (df = 3993)
F Statistic	3.988*** (df = 7; 6366)	17.154*** (df = 7; 16694)	7.368*** (df = 7; 4338)	4.797*** (df = 7; 3993)
Note:				*p**p***p<0.01

TABLE D.2. Interrupted Times Series Regression on Censorshipness Across Genre

Interrupted Time Series Regression on Censorshipness

	Dependent variable:			
	Censorshipness (%)			
	Hip-Hop	*Pop*	*Rock*	Folk
Date	0.007	-0.004	0.015***	-0.014**
	(0.007)	(0.003)	(0.005)	(0.006)
I-ROC	0.274	-1.155***	-1.143*	0.112
	(0.776)	(0.415)	(0.648)	(0.831)
day-Post-ROC	-0.007	0.011***	-0.017***	0.026***
	(0.008)	(0.004)	(0.006)	(0.008)
I-Censor	-1.798***	-0.816**	1.232*	-0.881
	(0.690)	(0.368)	(0.740)	(0.800)
day-Post-Censor	0.011*	-0.004	0.005	-0.004
	(0.006)	(0.003)	(0.007)	(0.007)
I-ROC2	0.189	-0.518	-0.008	0.545
	(0.644)	(0.379)	(0.762)	(0.866)
day-Post-ROC2	-0.019***	0.006*	-0.004	-0.031***
	(0.007)	(0.004)	(0.008)	(0.008)
Constant	62.445***	53.319***	53.507***	50.407***
	(0.733)	(0.348)	(0.595)	(0.663)
Observations	6,374	16,702	4,346	4,001
R^2	0.003	0.004	0.006	0.009
Adjusted R^2	0.001	0.003	0.004	0.008
Residual Std. Error	10.243 (df = 6366)	8.660 (df = 16694)	8.172 (df = 4338)	9.421 (df = 3993)
F Statistic	2.321** (df = 7; 6366)	9.288*** (df = 7; 16694)	3.760*** (df = 7; 4338)	5.330*** (df = 7; 3993)
Note:				*p**p***p<0.01

TABLE D.3. Interrupted Times Series Regression on Hip-Hopiness Across Genre and Popularity

Interrupted Time Series Regression on Hip-Hopiness

	Dependent variable:							
	Hip-Hopiness (%)							
	Hip-Hop		*Pop*		*Rock*		*Folk*	
	Top10%	Bottom50%	Top10%	Bottom50%	Top10%	Bottom50%	Top10%	Bottom50%
Date	0.019^{*}	-0.006	0.0002	-0.012^{***}	0.013	-0.001	-0.048^{***}	-0.023^{***}
	(0.010)	(0.008)	(0.006)	(0.004)	(0.012)	(0.006)	(0.015)	(0.008)
I-ROC	-0.486	1.208	-1.010	-0.200	-1.151	0.521	4.916^{***}	0.024
	(1.325)	(0.804)	(0.792)	(0.528)	(1.361)	(0.738)	(1.821)	(1.040)
day-Post-ROC	-0.014	0.002	0.008	0.014^{***}	-0.017	-0.010	0.046^{***}	0.030^{***}
	(0.012)	(0.009)	(0.007)	(0.005)	(0.015)	(0.007)	(0.017)	(0.010)
I-Censor	-2.125	-1.795^{***}	-1.383	-1.294^{***}	1.968	4.822^{***}	1.028	-0.702
	(1.388)	(0.674)	(0.843)	(0.403)	(1.804)	(0.801)	(2.291)	(0.917)
day-Post-Censor	0.013	0.026^{***}	-0.001	0.009^{**}	0.001	-0.014^{*}	-0.008	0.013
	(0.013)	(0.006)	(0.008)	(0.004)	(0.016)	(0.007)	(0.021)	(0.008)
I-ROC2	-0.700	-1.677^{***}	-0.505	-0.385	0.482	2.767^{***}	0.665	-0.270
	(1.707)	(0.572)	(0.876)	(0.418)	(1.500)	(0.813)	(2.643)	(0.932)
day-Post-ROC2	-0.050^{**}	-0.026^{***}	-0.012	-0.003	0.017	0.016^{*}	-0.004	-0.040^{***}
	(0.020)	(0.006)	(0.009)	(0.004)	(0.017)	(0.008)	(0.027)	(0.009)
Constant	64.024^{***}	68.282^{***}	58.982^{***}	60.405^{***}	57.451^{***}	59.278^{***}	59.746^{***}	57.570^{***}
	(1.051)	(0.901)	(0.623)	(0.468)	(1.333)	(0.661)	(1.364)	(0.859)
Observations	620	3,154	1,521	8,376	423	2,167	396	1,977
R^2	0.030	0.009	0.004	0.014	0.047	0.020	0.034	0.022
Adjusted R^2	0.019	0.007	-0.001	0.013	0.030	0.017	0.017	0.019
Residual Std. Error	6.566 (df = 612)	6.851 (df = 3146)	5.782 (df = 1513)	6.835 (df = 8368)	5.198 (df = 415)	6.431 (df = 2159)	7.158 (df = 388)	7.598 (df = 1969)
F Statistic	2.749^{***} (df = 7; 612)	4.158^{***} (df = 7; 3146)	0.829 (df = 7; 1513)	16.987^{***} (df = 7; 8368)	2.894^{***} (df = 7; 415)	6.416^{***} (df = 7; 2159)	1.973^{*} (df = 7; 388)	6.473^{***} (df = 7; 1969)
Note:								*p**p***p<0.01

TABLE D.4. Interrupted Times Series Regression on Censorshipness Across Genre and Popularity

Interrupted Time Series Regression on Censorshipness

	Dependent variable:							
	Censorshipness (%)							
	Hip-Hop		*Pop*		*Rock*		*Folk*	
	Top10%	Bottom50%	Top10%	Bottom50%	Top10%	Bottom50%	Top10%	Bottom50%
Date	0.021	-0.012	0.011	-0.010*	0.036**	0.015*	-0.009	-0.005
	(0.015)	(0.012)	(0.007)	(0.006)	(0.017)	(0.008)	(0.019)	(0.010)
I-ROC	-1.591	2.203*	-1.935**	-0.908	-4.994***	0.743	0.447	-1.455
	(2.028)	(1.226)	(0.963)	(0.739)	(1.924)	(0.994)	(2.331)	(1.280)
day-Post-ROC	-0.001	0.010	-0.004	0.020***	-0.042**	-0.026***	0.010	0.025**
	(0.019)	(0.013)	(0.009)	(0.007)	(0.021)	(0.010)	(0.022)	(0.012)
I-Censor	-6.588***	-1.848*	-0.852	-1.294**	-7.177***	3.624***	1.608	-1.637
	(2.125)	(1.028)	(1.025)	(0.563)	(2.550)	(1.078)	(2.932)	(1.128)
day-Post-Censor	0.023	0.015*	-0.0001	-0.004	0.066***	-0.001	0.007	-0.009
	(0.020)	(0.009)	(0.010)	(0.006)	(0.022)	(0.010)	(0.026)	(0.010)
I-ROC2	-3.251	-0.030	-1.746	-0.348	-3.275	0.997	-6.600*	0.421
	(2.613)	(0.872)	(1.065)	(0.585)	(2.119)	(1.095)	(3.383)	(1.146)
day-Post-ROC2	-0.042	-0.022**	0.002	0.001	-0.052**	0.012	0.001	-0.041***
	(0.030)	(0.009)	(0.011)	(0.005)	(0.023)	(0.011)	(0.035)	(0.011)
Constant	60.110***	64.778***	51.434***	53.945***	52.372***	52.559***	49.012***	49.468***
	(1.609)	(1.374)	(0.757)	(0.654)	(1.884)	(0.890)	(1.746)	(1.057)
Observations	620	3,154	1,521	8,376	423	2,167	396	1,977
R^2	0.029	0.004	0.005	0.004	0.076	0.015	0.024	0.019
Adjusted R^2	0.018	0.001	0.0002	0.004	0.061	0.012	0.007	0.015
Residual Std. Error	10.049 (df = 612)	10.447 (df = 3146)	7.033 (df = 1513)	9.561 (df = 8368)	7.346 (df = 415)	8.660 (df = 2159)	9.163 (df = 388)	9.347 (df = 1969)
F Statistic	2.637** (df = 7; 612)	1.640 (df = 7; 3146)	1.041 (df = 7; 1513)	5.368*** (df = 7; 8368)	4.893*** (df = 7; 415)	4.761*** (df = 7; 2159)	1.370 (df = 7; 388)	5.438*** (df = 7; 1969)
Note:								*p**p***p<0.01

Appendix E. Robustness Check

I used two ways to check the robustness of the statistical findings in the main article. The first way was to use a placebo test where I checked if a statistically significant shift in *Hip-Hopiness* and *Censorshipness* would take place at a time prior to the censorship. If a statistically significant shift could be found, it would challenge our finding that the shift in *Hip-Hopiness* and *Censorshipness* was caused by the censorship.

In the test, I assumed that the intervention happened at an earlier time (and hence a placebo intervention) and then ran a similar interrupted time series model as the one in the main article. Specifically, I used all the data in Period I and II (from Jan 1, 2017, to Jan 19, 2018) for the model, assuming there were two interventions within this period: one actual intervention (the broadcasting of *The Rap of China* on Jun 24, 2017) and one placebo (the assumed intervention). I arbitrarily chose Oct 20, 2017, as the time of the assumed intervention as it is at around the midpoint of the two actual interventions (the broadcasting of *The Rap of China* and the censorship) and there was no notable event at the time that might make a significant impact on how Hip-Hop musicians make music. I also excluded the data posterior to the actual censorship since it would confuse the modeling of the placebo test. The results of the test are shown in the tables below. As we can see, the placebo test suggests that the assumed intervention does not make any statistically significant impact on the *Hip-Hopiness* of songs in any of the four genres. The results hold even if we take popularity into consideration.

TABLE E.1. Interrupted Times Series Regression on Hip-Hopiness Across Genre, Placebo Test

Interrupted Time Series Regression on Hip-Hopiness, Placebo Test (10/20/2017)

	Dependent variable:			
	Hip-Hopiness (%)			
	Hip-Hop	*Pop*	*Rock*	*Folk*
Date	0.004	-0.007***	0.002	-0.016***
	(0.004)	(0.002)	(0.004)	(0.005)
I-ROC	0.141	-0.133	-1.127*	0.852
	(0.605)	(0.352)	(0.582)	(0.808)
day-Post-ROC	-0.006	0.010**	0.006	0.006
	(0.008)	(0.005)	(0.008)	(0.011)
Placebo	-0.132	0.665	-0.699	1.124
	(0.748)	(0.410)	(0.691)	(1.055)
day-Post-Placebo	0.002	-0.008	-0.011	0.015
	(0.012)	(0.007)	(0.012)	(0.018)
Constant	66.771***	59.676***	58.746***	56.763***
	(0.486)	(0.242)	(0.447)	(0.545)
Observations	2,626	7,101	2,459	1,960
R^2	0.001	0.003	0.005	0.009
Adjusted R^2	-0.001	0.003	0.002	0.006
Residual Std. Error	6.797 (df = 2620)	6.024 (df = 7095)	6.142 (df = 2453)	7.746 (df = 1954)
F Statistic	0.458 (df = 5; 2620)	4.762*** (df = 5; 7095)	2.224** (df = 5; 2453)	3.486*** (df = 5; 1954)
Note:				*p**p***p<0.01

TABLE E.2. Interrupted Times Series Regression on Hip-Hopiness Across Genre and Popularity, Placebo Test

Interrupted Time Series Regression on Hip-Hopiness Across Popularity, Placebo Test (10/20/2017)

	Dependent variable:							
	Hip-Hopiness (%)							
	Hip-Hop		*Pop*		*Rock*		*Folk*	
	Top10%	Bottom50%	Top10%	Bottom50%	Top10%	Bottom50%	Top10%	Bottom50%
Date	0.031**	-0.005	0.001	-0.010***	0.015	-0.001	-0.044***	-0.021**
	(0.013)	(0.007)	(0.007)	(0.003)	(0.012)	(0.006)	(0.016)	(0.008)
I-ROC	-0.650	1.027	-0.959	-0.042	-2.221	0.239	7.327***	0.552
	(2.036)	(0.905)	(1.030)	(0.546)	(1.705)	(0.852)	(2.489)	(1.199)
day-Post-ROC	-0.037	0.004	-0.004	0.007	-0.007	-0.002	-0.008	0.011
	(0.028)	(0.013)	(0.014)	(0.007)	(0.025)	(0.012)	(0.033)	(0.015)
Placebo	-2.437	0.094	1.342	0.980	1.600	-1.104	8.400	0.511
	(2.526)	(1.101)	(1.277)	(0.596)	(2.289)	(1.025)	(6.835)	(1.531)
day-Post-Placebo	0.075*	-0.011	0.014	-0.002	-0.055	0.002	-0.015	0.033
	(0.042)	(0.018)	(0.022)	(0.010)	(0.039)	(0.017)	(0.108)	(0.027)
Constant	62.378***	67.948***	59.108***	60.082***	57.325***	59.265***	58.547***	57.361***
	(1.361)	(0.830)	(0.705)	(0.375)	(1.257)	(0.651)	(1.440)	(0.852)
Observations	237	1,318	647	3,603	241	1,220	195	982
R^2	0.062	0.003	0.012	0.006	0.021	0.009	0.059	0.017
Adjusted R^2	0.041	-0.001	0.004	0.005	0.0001	0.005	0.034	0.012
Residual Std. Error	6.602 (df = 231)	7.123 (df = 1312)	5.708 (df = 641)	6.194 (df = 3597)	5.024 (df = 235)	6.551 (df = 1214)	6.556 (df = 189)	8.198 (df = 976)
F Statistic	3.033** (df = 5; 231)	0.665 (df = 5; 1312)	1.556 (df = 5; 641)	4.575*** (df = 5; 3597)	1.005 (df = 5; 235)	2.247** (df = 5; 1214)	2.363** (df = 5; 189)	3.345*** (df = 5; 976)
Note:								*p**p***p<0.01

The second way I tested the robustness of the results was to use a subset of the data to see if the results still hold. I randomly sampled 75% of the songs out of the dataset in proportion to the number of songs of different genres, popularity, and the period in which the song was released. This led to a sample of 27374 songs out of the total 36052 songs released in 2017 and 2018. I ran the same interrupted time series model as the one in the main article. The results, shown as follows, confirm that the impact of the censorship is still statistically significant even if we only use a subset of the data.

TABLE E.3. Interrupted Times Series Regression on Hip-Hopiness Across Genre, 75% sample

Interrupted Time Series Regression on Hip-Hopiness, 75% Sample

	Dependent variable:			
	Hip-Hopiness (%)			
	Hip-Hop	*Pop*	*Rock*	*Folk*
Date	0.007	-0.006**	0.001	-0.018***
	(0.005)	(0.003)	(0.005)	(0.006)
I-ROC	-0.432	-0.111	-0.147	0.327
	(0.599)	(0.355)	(0.554)	(0.785)
day-Post-ROC	-0.006	0.009***	-0.004	0.023***
	(0.006)	(0.003)	(0.006)	(0.007)
I-Censor	-1.492***	-1.173***	2.948***	-0.154
	(0.527)	(0.316)	(0.635)	(0.762)
day-Post-Censor	0.013***	0.007**	-0.010*	0.002
	(0.005)	(0.003)	(0.006)	(0.007)
I-ROC2	-0.426	-0.523	1.218*	0.832
	(0.492)	(0.327)	(0.651)	(0.824)
day-Post-ROC2	-0.021***	-0.003	0.010	-0.026***
	(0.005)	(0.003)	(0.007)	(0.008)
Constant	66.668***	59.495***	58.576***	56.765***
	(0.556)	(0.297)	(0.517)	(0.631)
Observations	4,783	12,506	3,248	3,003
R^2	0.004	0.008	0.013	0.010
Adjusted R^2	0.003	0.007	0.010	0.008
Residual Std. Error	6.778 (df = 4775)	6.425 (df = 12498)	6.070 (df = 3240)	7.747 (df = 2995)
F Statistic	3.021*** (df = 7; 4775)	14.108*** (df = 7; 12498)	5.863*** (df = 7; 3240)	4.384*** (df = 7; 2995)
Note:				*p**p***p<0.01

TABLE E.4. Interrupted Times Series Regression on Hip-Hopiness Across Genre and Popularity, 75% Sample

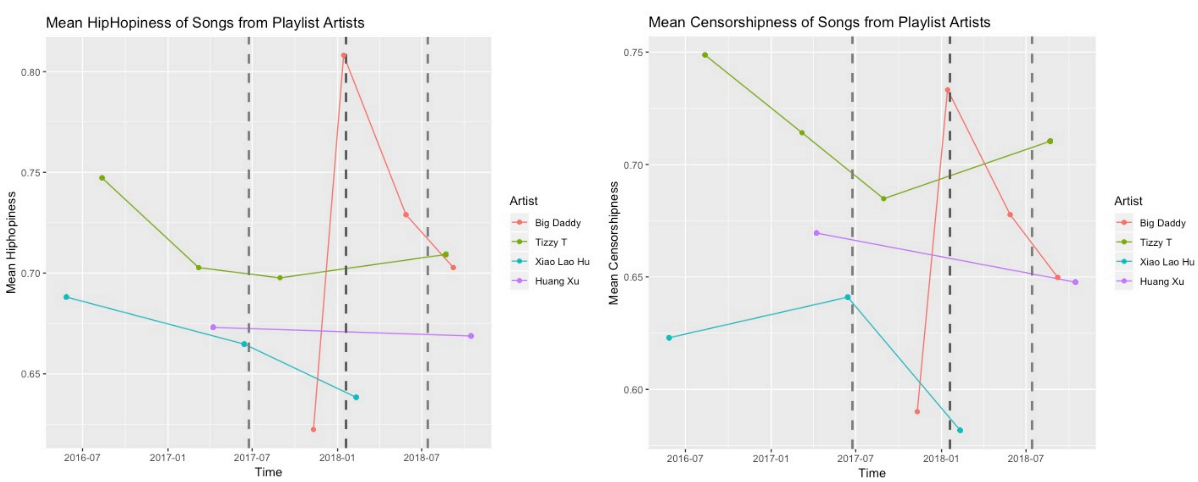
Interrupted Time Series Regression on Hip-Hopiness, 75% Sample

	Dependent variable:							
	Hip-Hopiness (%)							
	Hip-Hop		*Pop*		*Rock*		*Folk*	
	Top10%	Bottom50%	Top10%	Bottom50%	Top10%	Bottom50%	Top10%	Bottom50%
Date	0.012	-0.006	-0.001	-0.009*	0.006	-0.007	-0.050***	-0.020**
	(0.011)	(0.009)	(0.007)	(0.005)	(0.014)	(0.007)	(0.017)	(0.009)
I-ROC	-0.286	0.630	-0.268	-0.315	-0.445	1.597*	5.265**	-0.260
	(1.506)	(0.963)	(0.893)	(0.614)	(1.550)	(0.852)	(2.050)	(1.203)
day-Post-ROC	-0.006	0.004	0.006	0.011*	-0.009	-0.006	0.048**	0.027**
	(0.014)	(0.010)	(0.008)	(0.006)	(0.017)	(0.008)	(0.020)	(0.011)
I-Censor	-1.788	-1.902**	-1.014	-1.449***	2.285	4.726***	2.518	-0.524
	(1.581)	(0.789)	(0.952)	(0.469)	(2.133)	(0.919)	(2.708)	(1.056)
day-Post-Censor	0.003	0.026***	0.0004	0.013***	-0.002	-0.008	-0.025	0.014
	(0.015)	(0.007)	(0.009)	(0.005)	(0.018)	(0.008)	(0.025)	(0.010)
I-ROC2	1.376	-1.296*	-0.017	-0.792	0.629	2.247**	1.505	-0.280
	(1.962)	(0.671)	(0.988)	(0.486)	(1.814)	(0.942)	(3.034)	(1.076)
day-Post-ROC2	-0.053**	-0.031***	-0.011	-0.005	0.022	0.010	0.032	-0.043***
	(0.023)	(0.007)	(0.010)	(0.005)	(0.020)	(0.009)	(0.032)	(0.010)
Constant	65.316***	68.439***	58.997***	60.059***	57.971***	59.470***	59.765***	56.927***
	(1.178)	(1.052)	(0.698)	(0.532)	(1.527)	(0.776)	(1.596)	(0.976)
Observations	468	2,364	1,147	6,266	318	1,617	296	1,485
R^2	0.023	0.010	0.002	0.016	0.049	0.021	0.037	0.025
Adjusted R^2	0.008	0.007	-0.004	0.015	0.027	0.016	0.014	0.020
Residual Std. Error	6.541 (df = 460)	6.933 (df = 2356)	5.663 (df = 1139)	6.849 (df = 6258)	5.308 (df = 310)	6.445 (df = 1609)	7.181 (df = 288)	7.602 (df = 1477)
F Statistic	1.552 (df = 7; 460)	3.546*** (df = 7; 2356)	0.318 (df = 7; 1139)	14.841*** (df = 7; 6258)	2.260** (df = 7; 310)	4.860*** (df = 7; 1609)	1.592 (df = 7; 288)	5.417*** (df = 7; 1477)
Note:								*p**p***p<0.01

Appendix F. Examination on Individual Musicians

The findings in the main article are also validated by examinations of specific musicians and their works. For example, there are four Hip-Hop musicians whose songs have been censored but still had their new song released after the censorship: Big Daddy, Tizzy T, Xiao Lao Hu, and Huang Xu. Big Daddy is a musician from Taiwan and the rest three are based in mainland China. The following GRAPHIC F.1 shows the change of *Hip-Hopiness* and *Censorshipness* of their songs. Songs released before 2017 and still available on Chinesemusic.com are also included for comparison. Means are taken if the musician released multiple songs on the same day. As GRAPHIC F.1 shows, the two musicians who released new songs in Period III, Big Daddy and Xiao Lao Hu (from mainland China), have decreased both *Hip-Hopness* and *Censorshipness* in their song released in Period 3 compared to their previous songs. Both Tizzy T and Huang Xu did not release new songs in Period III but in Period IV. While Tizzy slightly increases both *Hip-Hopness* and *Censorshipness* in his new songs released in Period IV compared to his previous release in Period 2, Huang Xu basically maintains the *Hip-Hopness* and only slightly decreases *Censorshipness* compared to his previous release before 2017. It is also worth noting that Huang Xu changed his major genre from Hip-Hop before 2018 to Pop in his new release in Period IV, which indicates one way Hip-Hop musicians manage to “circumvent” the censorship while still keeping doing similar kind of music.

GRAPHIC F.1. Hip-Hopiness and Censorshipness of Songs by Musicians Whose Songs were Censored by the Censorship



Bibliography

Askin, Noah, and Michael Mauskapf. 2017. "What Makes Popular Culture Popular? Product Features and Optimal Differentiation in Music." *American Sociological Review* 82(5):910–944.

Baranovitch, Nimrod. 2003. *China's New Voices: Popular Music, Ethnicity, Gender, and Politics, 1978-1997*. Univ of California Press.

DiMaggio, Paul, Manish Nag, and David Blei. 2013. "Exploiting Affinities between Topic Modeling and the Sociological Perspective on Culture: Application to Newspaper Coverage of US Government Arts Funding." *Poetics* 41(6):570–606.

Hamel, Philippe, and Douglas Eck. 2010. "Learning Features from Music Audio with Deep Belief Networks." Pp. 339–344 in *ISMIR*. Vol. 10. Utrecht, The Netherlands.

Hannan, Michael T. 2010. "Partiality of Memberships in Categories and Audiences." *Annual Review of Sociology* 36(1):159–81. doi: 10.1146/annurev-soc-021610-092336.

Hannan, Michael T., Michael T. Hannan, László Pólos, and Glenn R. Carroll. 2007. *Logics of Organization Theory: Audiences, Codes, and Ecologies*. Princeton University Press.

Long, J. Scott, and Laurie H. Ervin. 2000. "Using Heteroscedasticity Consistent Standard Errors in the Linear Regression Model." *The American Statistician* 54(3):217–224.

MacKinnon, James G., and Halbert White. 1985. "Some Heteroskedasticity-Consistent Covariance Matrix Estimators with Improved Finite Sample Properties." *Journal of Econometrics* 29(3):305–325.

McFee, Brian, Colin Raffel, Dawen Liang, Daniel PW Ellis, Matt McVicar, Eric Battenberg, and Oriol Nieto. 2015. "Librosa: Audio and Music Signal Analysis in Python." in *Proceedings of the 14th python in science conference*. Vol. 8.

Roberts, Margaret E., Brandon M. Stewart, and Dustin Tingley. 2014. "Stm: R Package for Structural Topic Models." *Journal of Statistical Software* 10(2):1–40.

Trohidis, Konstantinos, Grigorios Tsoumakas, George Kalliris, and Ioannis P. Vlahavas. 2008. "Multi-Label Classification of Music into Emotions." Pp. 325–330 in *ISMIR*. Vol. 8.

Xu, Changsheng, Namunu Chinthaka Maddage, and Xi Shao. 2005. "Automatic Music Classification and Summarization." *IEEE Transactions on Speech and Audio Processing* 13(3):441–450.