Normalization of Censorship: Evidence from China
Online Appendices

Tony Zirui Yang

March 31, 2021

Contents

Appendices

Appendix A Survey Procedure and Descriptive Statistics

A.1 Survey Procedure

The survey was administered in mainland China by a Shanghai-based Chinese online survey company *Kurun Data* (库润数据: `https://www.kurundata.com/`). The 612 participants were recruited by *Kurun* and then directed to a US-based website, Qualtrics, where they completed the survey anonymously. Once they completed the survey on Qualtrics, they were redirected back to *Kurun*'s platform.

All mainland Chinese citizens above 18 years old are eligible for this study. To make sure that the sample covers a broad range of socioeconomic backgrounds, I put quotas on gender, education, and age. In the end, the gender quota successfully yields a gender-balanced sample. The education and age quotas alleviate the problem of homogeneous survey participants but fall short of yielding a sample representative of the Internet population.

A.2 Compliance with Ethical Principles of Human Subject Research

The survey followed all established principles of human subject research and was approved by the Institutional Review Board (IRB) at the researcher's home institution. Although the IRB exempted this study from a formal consent form, I still included a consent page and information sheet at the beginning of the survey. All participants were informed about the purpose, the length, and the format of the study. All participants need to click "I consent" on the information sheet page before they could proceed. They were allowed to opt-out of the study at any point of the survey. Incomplete survey responses were not recorded.

Because the treatment prompt explicitly asked the respondents to **imagine** that they were reading WeChat articles, no deception was used. All articles in the experiment were actual WeChat articles that were censored by WeChat. At the end of the survey, participants were explicitly told that this is an experimental study and information in the survey might not be representative of the reality.

All respondents were paid by *Kurun* at its usual rate for their participation. *Kurun* was paid by the researcher of this study. All participants were adults and none of them would be put in a disadvantageous position had they chosen not to participate.

Because this survey was conducted in China, an authoritarian regime, I paid extra caution to protect respondents' information and responses, so that they will not be negatively affected by the authority due to their participation in this study. I did not ask for personal information that could directly identify participants' identity, such as names, phone numbers, and email addresses. I stored all the responses at Qualtrics via an American institutional account. The study passed the information security review at the researcher's home institution.

A.3 Summary Statistics and Demographics

Table A1: Summary Statistics

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Treatment	612	0.498	0.500	0	0	1	1
Female	608	0.500	0.500	0	0	1	1
Age	611	3.897	1.570	1	3	5	8
Education	611	3.427	0.890	1	3	4	5
Income	605	3.231	1.101	1	3	4	5
Party Member	610	0.261	0.439	0	0	1	1
Pro-Market Ideology	610	2.605	1.307	1	1	4	5
Political Interests	611	4.103	1.319	1	3	5	6
Social Media	611	3.355	1.026	1	3	4	5
VPN Usage	609	2.125	0.867	1	2	3	4
Support for Control	603	3.484	1.019	1	3	4	5
Censorship is Normal	607	3.802	0.941	1	3	4	5
Perceived Volume	606	2.888	0.816	1	2	3	5
Central Government	610	4.075	0.886	1	4	5	5
Local Government	597	3.829	0.990	1	3	5	5
Overall Satisfaction	611	3.974	0.910	1	4	5	5
Willing to Protest	609	2.762	1.261	1	2	4	5

Note: VPN Usage is measured post-treatment due to concern of contaminating the treatment. Because it asks past experiences, it is unlikely to be affected by the treatment.

Table A2: Sociodemographics of the Study Participants and Chinese Internet Users

	Sociodemographics	Study Participant	Chinese Internet Users
Location	Urban	71.9%	71.8%
	Rural	28.1%	28.2%
Region	Eastern China	50.8%	41.2%
	Northeastern China	7.8%	8.4%
	Central China	19.6%	22.1%
	Western China	21.6%	23.3%
Occupation	Student	8.3%	26.9%
	Self-employed	13.1%	22.4%
	Corporate office worker	34.5%	8.0%
	Corporate management	16.3%	2.9%
	Government employee	2.8%	2.8%
	Professional	12.6%	6.0%
	Manufacturing worker	4.2%	2.6%
	Service worker	3.6%	4.4%
	Migrant worker	2.0%	4.2%
	Farmer	0.7%	6.3%
	Unemployed	0.7%	8.8%
	Retired and other	1.3%	4.7%
Gender	Female	49.7%	48.1%
	Male	49.7%	51.9%
Education	Junior high school or lower	3.6%	56.1%
	Senior high school	12.6%	23.8%
	3-year college	25.2%	10.5%
	4-year college or higher	58.5%	9.7%
Age	≤ 19	6.4%	23.2%
	20-29	31.4%	21.5%
	30-39	45.1%	20.8%
	40-49	14.9%	17.6%
	≥ 50	2.1%	16.9%
Income	≤ 3000	7.8%	51.0%
	3000-5000	13.9%	21.5%
	5000-8000	38.2%	14.3%
	≥ 8000	38.9%	13.3%

Note: Data about Chinese Internet users are from *The 45th Statistical Report of Internet Development in China*, issued by China Internet Network Information Center (CNNIC) in April 2020. The education distribution of Chinese Internet users is taken from the 44th report of June 2019. The geographic distribution is taken from the 37th report of January 2016.

A.4 Balance Table

Table A3: Balance Table

	Obs.	Control Group	Treatment Group	p-value
Age	611	3.863	3.931	0.591
Income	605	3.221	3.242	0.818
Female	608	0.469	0.531	0.124
Education	611	3.407	3.447	0.577
Party Member	610	0.247	0.274	0.459
Political Interests	611	4.121	4.085	0.738
Pro-Market Ideology	610	2.668	2.541	0.232
Social Media Usage	611	3.313	3.398	0.304
VPN Usage	609	2.151	2.099	0.458

Table A4: Using Covariates to Predict Treatment

	Dependent variable:
	Treatment
Female	0.047
	(0.042)
Education	0.018
	(0.028)
Age	0.014
	(0.015)
Income	0.003
	(0.024)
Pro-Market Ideology	−0.011
	(0.016)
Party Member	0.035
	(0.050)
Political Interests	−0.019
	(0.019)
Social Media Usage	0.023
	(0.021)
VPN Usage	−0.013
	(0.025)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table A3 shows that all nine covariates are balanced between the treatment and control groups. Table A4 shows that none of the covariates significantly predicts treatment assignment. Hence, the randomization is successful.

Appendix B Regression Analysis

B.1 Full Results of Tables in the Main Paper

Table B1: Regression Analyses of Treatment Effect on Policy Backlash

	Dependent variable:		
	Support	Normal	Volume
	(1)	(2)	(3)
Treatment	0.205***	0.199***	−0.084
	(0.075)	(0.068)	(0.067)
Female	0.116	−0.017	−0.068
	(0.077)	(0.069)	(0.069)
Education	−0.013	0.102**	0.016
	(0.051)	(0.045)	(0.045)
Age	0.015	0.013	0.027
	(0.028)	(0.025)	(0.025)
Income	0.104**	0.007	0.030
	(0.044)	(0.039)	(0.039)
Pro-Market Ideology	−0.323***	−0.297***	0.030
	(0.029)	(0.026)	(0.026)
Party Member	0.121	0.288***	−0.034
	(0.091)	(0.082)	(0.081)
Political Interests	−0.034	−0.007	0.004
	(0.034)	(0.030)	(0.030)
Social Media Usage	−0.012	0.046	−0.026
	(0.038)	(0.034)	(0.033)
VPN Usage	−0.072	−0.065	0.054
	(0.045)	(0.041)	(0.040)
Constant	4.123***	3.999***	2.581***
	(0.254)	(0.229)	(0.225)
Observations	581	585	585
R^2	0.220	0.255	0.020
Adjusted R^2	0.207	0.242	0.003
Note:			*p<0.1; **p<0.05; ***p<0.01

Table B2: Regression Analyses of Treatment Effect on Regime Backlash

	Dependent variable:			
	Overall	Central	Local	Collective
	(1)	(2)	(3)	(4)
Treatment	0.181***	0.201***	0.234***	−0.284***
	(0.068)	(0.067)	(0.076)	(0.102)
Female	0.071	−0.040	−0.010	−0.084
	(0.069)	(0.069)	(0.078)	(0.105)
Education	0.054	0.066	0.029	−0.019
	(0.046)	(0.045)	(0.051)	(0.069)
Age	0.012	0.002	0.026	0.073*
	(0.025)	(0.025)	(0.028)	(0.038)
Income	0.012	0.022	−0.050	−0.028
	(0.039)	(0.039)	(0.044)	(0.059)
Pro-Market Ideology	−0.211***	−0.192***	−0.248***	−0.069*
	(0.027)	(0.026)	(0.030)	(0.040)
Party Member	0.288***	0.192**	0.236**	0.367***
	(0.082)	(0.081)	(0.092)	(0.124)
Political Interests	0.071**	0.045	0.039	−0.122***
	(0.031)	(0.030)	(0.034)	(0.046)
Social Media Usage	−0.028	−0.029	−0.014	−0.039
	(0.034)	(0.033)	(0.038)	(0.051)
VPN Usage	−0.228***	−0.172***	−0.182***	0.216***
	(0.041)	(0.040)	(0.046)	(0.061)
Constant	4.340***	4.431***	4.540***	3.051***
	(0.229)	(0.226)	(0.256)	(0.344)
Observations	590	588	577	587
R^2	0.201	0.161	0.171	0.065
Adjusted R^2	0.187	0.147	0.156	0.049
Note:				*p<0.1; **p<0.05; ***p<0.01

B.2 Aggregate Results

In the Pre-Analysis Plan, I treat the main outcome variables, *Policy Backlash* and *Regime Backlash*, as aggregate measures combining those survey questions in section 2. However, it turns out that the internal consistency among survey items is low ($\alpha = 0.59$ for policy questions, and 0.56 for regime questions). Hence, it does not satisfy the assumption of aggregating survey items and I only report treatment effects on individual items in the main paper. Nevertheless, I combine the survey items by simple addition and re-scale them to vary from 0 to 1. Table B3 reports the treatment effect on aggregate measures. The results still hold.

Table B3: Treatment Effect on Aggregate Policy and Regime Backlash

	Dependent variable:			
	Policy Backlash		Regime Backlash	
Treatment	-0.048^{***}	-0.039^{***}	-0.069^{***}	-0.060^{***}
	(0.014)	(0.013)	(0.014)	(0.014)
Female		-0.015		-0.006
		(0.013)		(0.014)
Education		-0.007		-0.009
		(0.009)		(0.009)
Age		0.001		0.002
		(0.005)		(0.005)
Income		-0.007		-0.001
		(0.007)		(0.008)
Pro-Market Ideology		0.055^{***}		0.038^{***}
		(0.005)		(0.005)
Party Member		-0.037^{**}		-0.020
		(0.015)		(0.016)
Political Interests		0.004		-0.019^{***}
		(0.006)		(0.006)
Social Media		-0.005		0.001
		(0.006)		(0.007)
VPN Usage		0.017^{**}		0.054^{***}
		(0.008)		(0.008)
Constant	0.406^{***}	0.284^{***}	0.357^{***}	0.245^{***}
	(0.010)	(0.043)	(0.010)	(0.045)
Observations	594	573	591	572
Adjusted R^2	0.018	0.221	0.036	0.190
Note:				$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

B.3 Heterogeneous Treatment Effect

Figure B1 and B2 show that the treatment effects are generally consistent across different subgroups. Treatment effects among pro-market respondents are generally higher than pro-state respondents. This might be due to ceiling effects among pro-state respondents. Further analyses show that there is not a significant interaction effect between treatment and the demographic covariates. Therefore, although the sample in the study does not reflect the general Internet population, it is unlikely that such unrepresentative sample causes bias.

Figure B1: Heterogeneous Treatment Effect on Policy Backlash

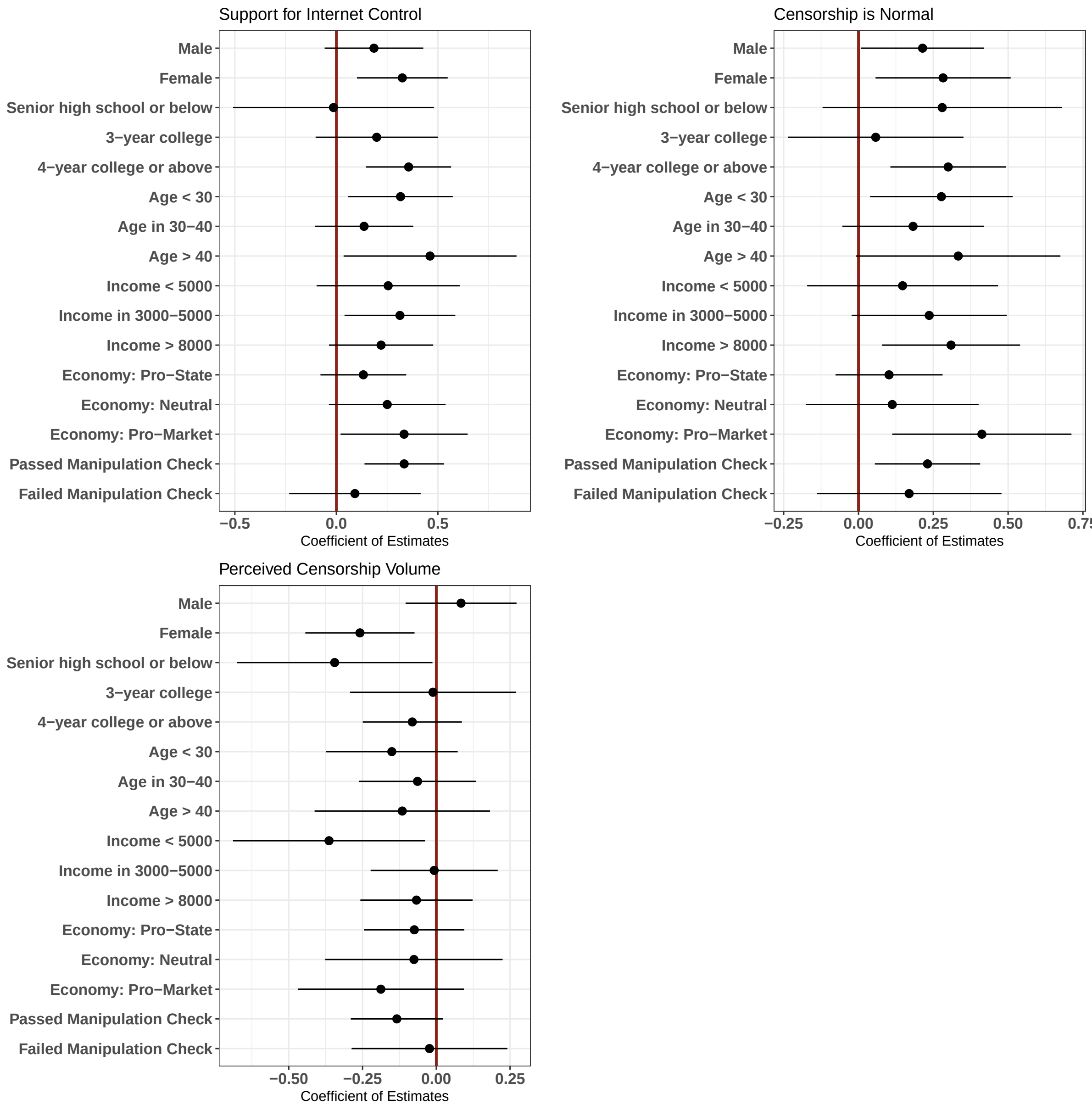
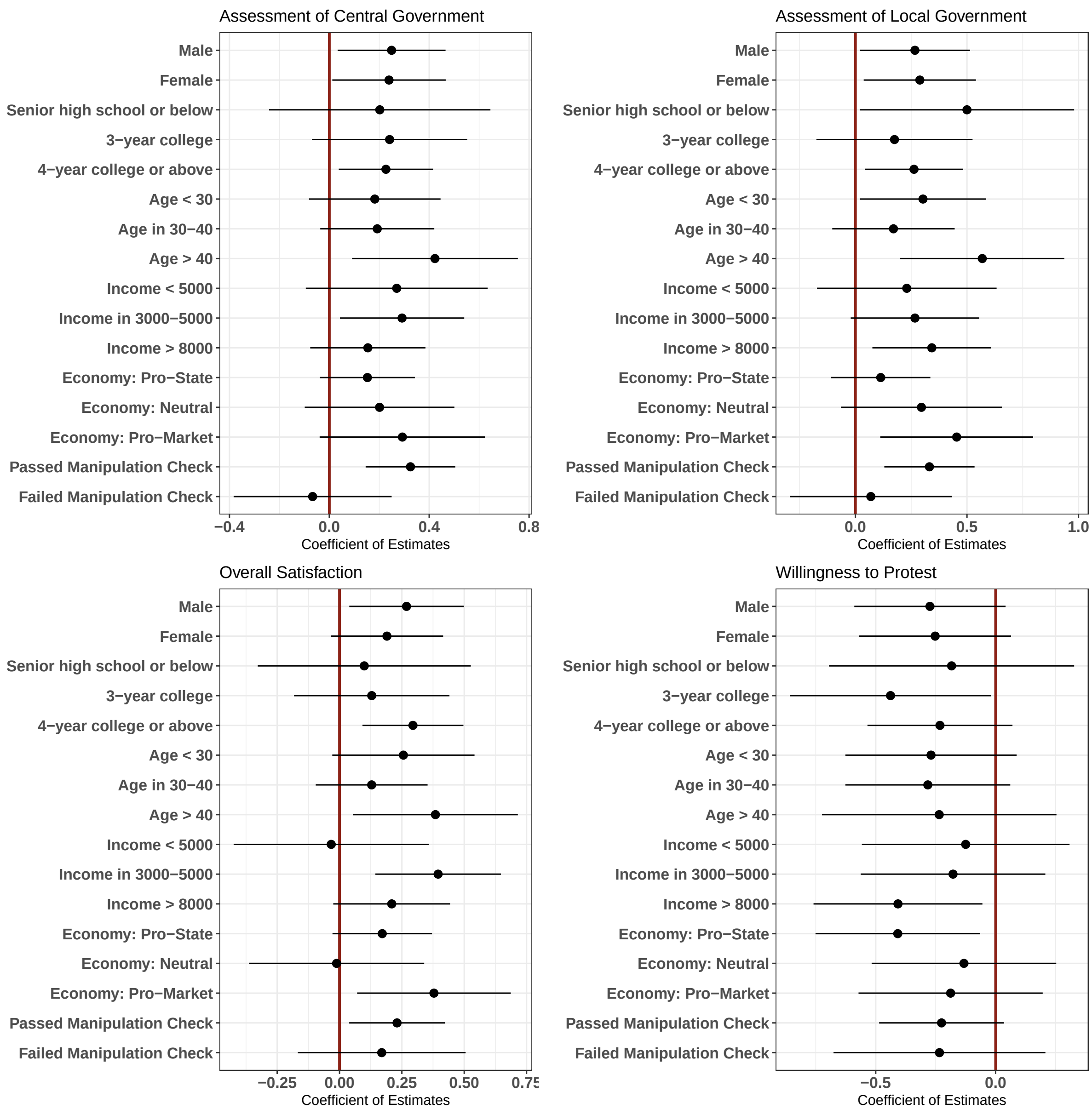


Figure B2: Heterogeneous Treatment Effect on Regime Backlash



As a manipulation check, respondents were asked to identify whether some specific article titles were labeled as censored by WeChat. Around 71% of the respondents passed the manipulation check. Respondents who passed manipulation check display larger and more significant treatment effects (except for *Willingness to Protest*), whereas respondents who failed manipulation check display insignificant treatment effects.

Appendix C Experiment Articles

Table C1: Treatment Articles (Order Randomized)

#	Collective Action or Government	Control Group	Treatment Group	Title
1	No		Censorship Label	银行哭晕了，一招教你多赚23倍活期高收益 The banks are crying, one trick to help you earn 23 times higher profit by demand deposit.
2	No		Censorship Label	男人究竟有多在意你的脸 How much do men care about your face.
3	No		Censorship Label	药王孙思邈活了142岁，临死前叮嘱弟子：务必将这个药方毁掉！ The King of traditional medicine Sun Simiao lived 142 years. Before he died, he told his pupils: Be sure to destroy this prescription!
4	No			关于华为离职员工被诉敲诈的十点疑问 Ten questions about Huawei's former employees being sued for extortion
5	No			告别996现状，创业艰难的我们将如何解压 Say goodbye to the stressful status quo. How can we relax ourselves under pressure of work.
6	No			他信英拉回梅州祭祖 Thaksin and Yingluck returned to Meizhou to worship ancestors.
7	Yes	Censorship Label	Censorship Label	刚刚，坪山佳士工人斗争取得初步胜利！ Just now, the Pingshan Jasic labor strike has won an initial victory!
8	Yes	Censorship Label	Censorship Label	江西"殡葬改革"切莫砸了棺材，伤了民心 Jiangxi's "Funeral Reform" must not smash the coffin and hurt people's hearts
9	Yes	Censorship Label	Censorship Label	胡温老了，那十年发生了什么 President Hu and Premier Wen are getting old. What happened in their decade.
10	Yes			个税调整后，你的收入下降了吗? After the tax reform, has your income decreased?

Appendix D Categorization of Censored Articles

D.1 Categories and Coding Process

In total, I keep track of nine different topic categories. In addition to three political categories: (1) collective action, (2) government criticism, and (3) government (non-criticism), I also include six non-political categories: (1) business & economy, (2) entertainment & sexuality, (3) advertisement, (4) local events, traditions, cultures, (5) foreign news, and (6) others.

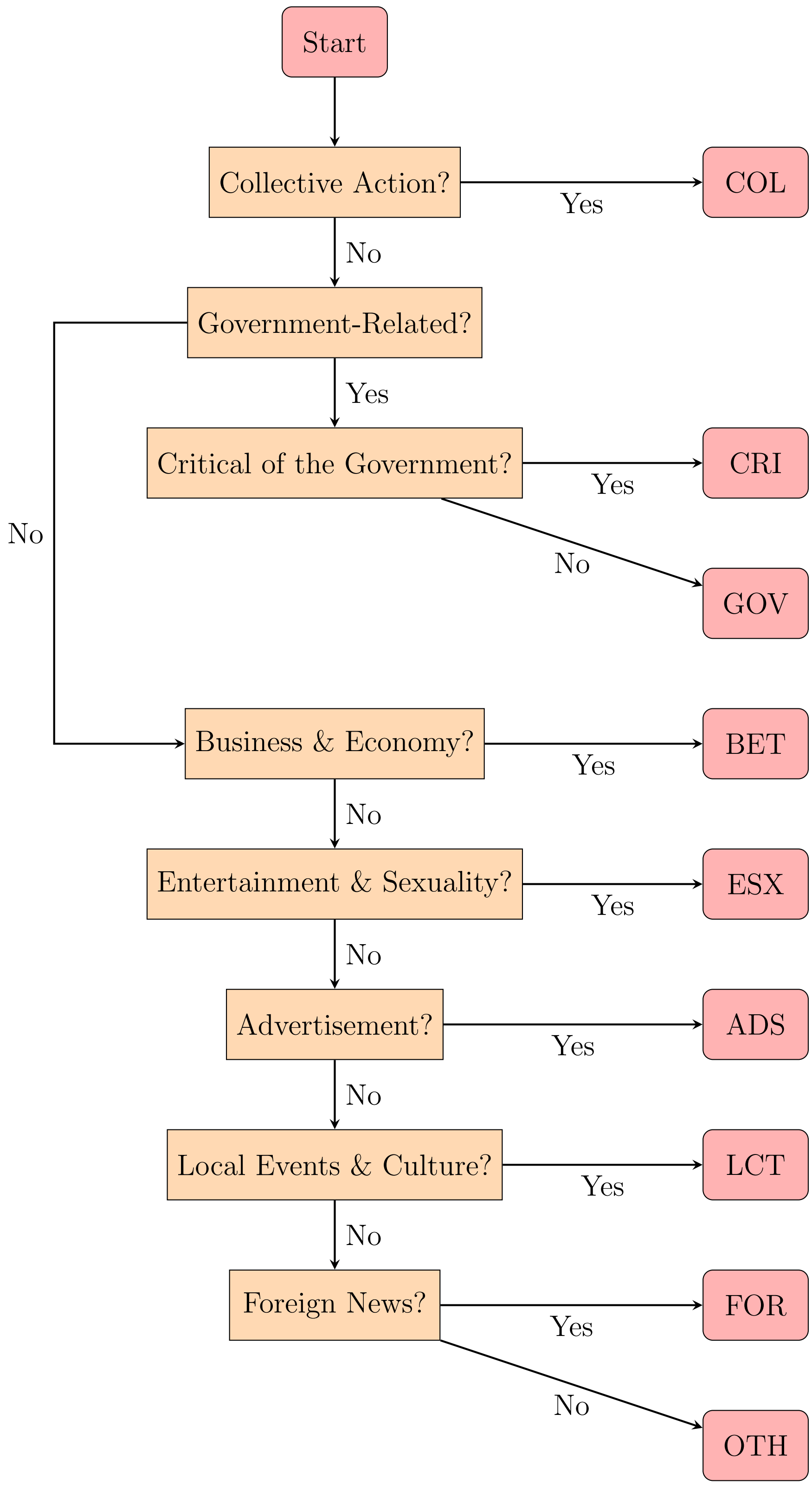
The categorization process mostly follows Miller (2018), but I also consider the categorization in King, Pan, and Roberts (2013). Although King, Pan, and Roberts (2013) is currently one of the most influential paper on censorship, Miller (2018) provides the most detailed and up-to-date categorization of censored content in China. In particular, the definition of collective action, business & economy, and entertainment & sexuality are the same as Miller (2018).

The definition of government-related content (i.e. government criticism and government non-criticism) combines the definitions of seven different categories in Miller (2018): *government*, *corruption*, *sensitive anniversary*, *recurring political event*, *regular political event*, *nationalism*, and *HK/Macau/Taiwan*. By using a broader definition of government-related content, I aim to establish the upper bound on the proportion of political content and avoid underestimating government-related content. The difference between government criticism and government non-criticism also follows the definition of government criticism in Miller (2018). Any government-related content that does not meet the definition of government criticism is categorized as government non-criticism.

The last four categories are created by myself due to the incompleteness of Miller's coding rules to my data. They are all self-explanatory. Importantly, none of these categories include politically salient events or issues. The last category is the residual category that includes all articles that do not fit into the definitions of the other groups.

One important difference from Miller (2018) is that the nine categories are mutually exclusive. Similar strategy is employed by King, Pan, and Roberts (2013). Having mutually exclusive categories simplifies the categorization process as well as the interpretation of the results. In practice, the nine categories are coded in sequence. An article will first be considered if it satisfies the definition of the collective action category. If yes, then the categorization process ends. If not, the article will then be considered if it belongs to the government criticism category and so on. If an article does not fit into the definitions of the first eight categories, it will be put into the last residual category. Such coding process ensures that collective action and government criticism will not be underestimated in the analysis. Figure D1 presents a flowchart of the coding process.

Figure D1: Flowchart of the Coding Process



D.2 Inter-Coder Reliability

Two coders coded the 2,500 articles in the training set independently. To code the training set, they both analyze the titles, the authors, and the content of the articles according to the coding rubric. Both coders are native Chinese graduate students in political science. Table D1 shows that their results are generally consistent in terms of the proportion of each topic category. The greatest disagreement between the two coders is whether an article belongs to Government Criticism (CRI) or Government Non-Criticism (GOV), which is not the main focus of this paper (also shown in table D3).

Table D2 shows that, when identifying whether an article is political or non-political, the two coders agree on 93.28% ($\frac{716+1616}{2500} = 0.9328$) of the cases. Table D3 shows the details of the two coders' coding. The accuracy rate between the two coders is 82.28% when considering specific topic categories. The macro F1 is 0.82 and the Cohen's κ between the two coders is 0.80, higher than the commonly applied criteria of 0.70 for inter-coder reliability tests.

Table D1: Two Coders' Coding of the Training Set of 2,500 Articles

	ADS	BET	COL	CRI	ESX	FOR	GOV	LCT	OTH	Total
Coder 1	250	345	69	422	440	167	311	311	185	2500
	10.0%	13.8%	2.8%	16.9%	17.6%	6.7%	12.4%	12.4%	7.4%	100%
Coder 2	237	344	72	309	432	141	417	300	248	2500
	9.5%	13.8%	2.9%	12.4%	17.3%	5.6%	16.7%	12.0%	9.9%	100%

Note: ADS: Advertisement. BET: Business & Economy. COL: Collective Action. CRI: Government Criticism. ESX: Entertainment & Sexuality. FOR: Foreign News. GOV: Government (non-criticism). LCT: Local Events, Culture, Traditions. OTH: Others.

Table D2: Inter-Coder Reliability (General Category)

		Coder 2		
		Political	Non-Political	Total
	Political	716	86	802
Coder 1	Non-Political	82	1616	1698
	Total	798	1702	2500

Note: Political content includes collective action (COL), Government Criticism (CRI), and Government Non-Criticism (GOV). Other categories are non-political content.

Table D3: Inter-Coder Reliability (Specific Category)

		Coder 2									
		ADS	BET	COL	CRI	ESX	FOR	GOV	LCT	OTH	
Coder 1	ADS	215	0	0	0	9	0	0	7	19	
	BET	4	306	1	2	8	5	6	4	9	
	COL	0	0	62	1	2	1	2	1	0	
	CRI	0	9	2	295	5	4	82	12	13	
	ESX	8	6	1	3	393	0	5	7	17	
	FOR	0	3	1	1	4	130	16	3	9	
	GOV	1	10	0	1	3	0	271	17	8	
	LCT	5	7	1	4	7	1	16	241	29	
	OTH	4	3	4	2	1	0	19	8	144	
		ADS	BET	COL	CRI	ESX	FOR	GOV	LCT	OTH	Macro
	Precision	0.91	0.89	0.86	0.95	0.91	0.92	0.65	0.80	0.58	0.83
	Recall	0.86	0.89	0.90	0.70	0.89	0.78	0.87	0.77	0.78	0.83
	F1	0.88	0.89	0.88	0.81	0.90	0.84	0.74	0.79	0.67	0.82

Note: Rows are coder 1 and columns are coder 2. ADS: Advertisement. BET: Business & Economy. COL: Collective Action. CRI: Government Criticism. ESX: Entertainment & Sexuality. FOR: Foreign News. GOV: Government (non-criticism). LCT: Local Events, Culture, Traditions. OTH: Others.

D.3 Example Articles

1. Collective Actions (COL)
   (a) 写入国务院文件的要求，群众拉横幅抗议！舆论哗然后教育部回应了
   The request written in the documents of the State Council; the masses raised banners to protest! Public opinion was upset and the Ministry of Education responded.
   (b) 刚刚，坪山佳士工人斗争取得初步胜利！
   Just now, the Pingshan Jiashi workers' struggle has won an initial victory!
   (c) 香港暴徒冲击手册曝光！密谋7·21漆弹袭警
   Hong Kong thugs' manual exposed! Plotting 7.21 Paintball Attack on Police.
   (d) 澳媒：香港暴力抗议者统治街头
   Australian media: Hong Kong violent protesters rule the streets.
   (e) 山东上访村民引爆炸药,与镇政府大院同归于尽你怎么看？欢迎转发！
   The petitioning villagers in Shandong Province detonated explosives, attempting to employ suicide bomb strategy at the township government. What do you think? Welcome to repost!
2. Government Criticism (CRI)
   (a) 武汉封城第43天，我问了方方一些日记之外的批评问题
   On the 43rd day of Wuhan's lockdown, I asked Fang Fang some critical questions outside of her diary.
   (b) 河南又火了，官员组团贪腐
   Henan province is under the spotlight again: collective corruption of the officials.
   (c) 征二胎税？！历史将永远记住这两个递刀的人！
   Second-child tax?! History will always remember these two complicit!
   (d) 任志强国务院开炮，大胆发言
   Ren Zhiqiang blasts the State Council.
   (e) 41人任命被撤销，违规提拔后，被退回官员怎么办？
   41 appointments were revoked. What should we do with illegally promoted officials?
3. Government Non-Criticism (GOV)
   (a) 中国四中全会期待改革再出发，应对大变局政策需落地
   The Fourth Plenary Session of China's Central Committee looks forward to reforms, and policies to respond to major changes need to be implemented.
   (b) 一文让你看懂中国特色社会主义制度的优越性！
   This article allows you to understand the superiority of the socialist system with Chinese characteristics!
   (c) 我国动真格了！直接取消德国千亿合同，令其肠子悔青，疾呼：有话好好说！
   Our country is showing strength! Cancelling the 100 billion contract with Germany. The Germans are in regret!
   (d) 台湾政坛大乱斗，大陆随时可能解放台湾！
   The political chaos in Taiwan. Mainland may liberate Taiwan anytime!

(e) 好消息，中国！即将实现打电话不用钱，上网不用流量！
Good news, China! It will soon be realized that cellphone and Internet data are free!

4. Business & Economy (BET)
   (a) 重磅!碧桂园再爆危机涉中信诚信窝案20亿利益输送或解地产界陈年谜案
   Big News! Another crisis of the Country Garden Group, involving the 2 billion CITIC case.
   (b) 谷歌或推出中国版搜索引擎茅台，百度股价闻声大跌近8%！
   Google may launch the Chinese version of the search engine: "Maotai". Baidu's stock price plummeted by nearly 8%!
   (c) 陈浩武：杨小凯经济学思想的时代意义（下）
   Chen Haowu: The Significance of Yang Xiaokai's Economic Thoughts (Part 2)
   (d) 港股复盘：恒指两天盘中变脸，再现暴跌70%股票
   Hong Kong stocks review: the Hang Seng Index changed course in two consecutive days, and the stocks plummeted 70% again
   (e) 滴滴总裁柳青发了封道歉信，湖畔大学学员：加油
   Liu Qing, CEO of Didi, issued a letter of apology.

5. Entertainment & Sexuality (ESX)
   (a) 男人究竟有多在意你的脸?
   How much do men care about your face?
   (b) 香港绝色女艳星：喂春药令李小龙命丧闺床，死后大拍激情片...
   Hong Kong's stunning female porn star: Bruce Lee was killed by feeding aphrodisiac, making a passionate movie after his death...
   (c) 台湾女教授深入讲解男女关系，太精彩了！不浪费一秒钟从头笑到尾
   Taiwanese female professors explain in depth the relationship between men and women: it is wonderful! Don't waste a second laughing from start to finish
   (d) 贾乃亮助理点赞李小璐不要脸？刘亦菲曾被劈腿？范冰冰姐弟关系揭秘?
   Jia Nailiang liked Li Xiaolu as shameless? Liu Yifei was cheated? Fan Bingbing's relationship between sister and brother revealed?
   (e) 48岁的闫妮惊艳春晚：瘦了30斤之后，气质简直比20岁的女儿还好！
   48-year-old Yan Ni was amazing at the Spring Festival Gala: After losing 30 pounds, her temperament is better than her 20-year-old daughter!

6. Advertisement (ADS)
   (a) 3斤枸杞不如1两它，最强天然青霉素，每天喝一杯，抵抗力强10倍！
   3 pounds of wolfberry is not as good as 1 tael of the natural penicillin. Drink a cup per day to strengthen your immunity 10 times!
   (b) 健康—经期来了可别再喝红糖水了，这样做排出体内十年垃圾
   Health—Don't drink brown sugar water anymore when your menstrual period is coming. Do this to get rid of the garbage in your body for ten years
   (c) 【强烈推荐】这东西补血养颜，女人要多吃，40岁像20岁！
   Strongly recommended! This food nourishes blood and beauty. Women should eat more. Makes you look like 20 years old when you are 40!

(d) 【推荐】原来明星大V们喜欢关注这些公众号
Recommended! It turns out that celebrities like to pay attention to these public accounts

(e) 第一次创业，我建议你选择小程序代理W
For the first time starting a business, I suggest you choose a mini program agent

7. Local Events, Culture, & Traditions (LCT)

(a) 故宫的规矩
Rules of the Forbidden City.

(b) 辣眼睛！幼儿园开学典礼竟大跳钢管舞！园长被开除...
OMG! Pole dance in kindergarten! The principal was fired...

(c) 难怪民国出大师，民国老课本美到心醉。现在的孩子都该看看
No wonder there were many masters during the Republic era. These old textbooks from the Republic era are so beautiful. Children should read them now.

(d) 6 Breakfast Items You Must Try in Beijing

(e) 许倬云：中国文化让我伤心的地方
Xu Zhuoyun: What makes me sad about Chinese culture.

8. Foreign News (FOR)

(a) 特朗普遭美国众议院弹劾，真能弹掉吗?
Trump was impeached by the US House of Representatives, but can he really be impeached?

(b) 德国的吸星大法：元首当年用铁和血没做到的事情，被你们用欧元做到了！
The German method of attracting money: What Hitler failed to do with iron and blood, was done by the Germans with euros!

(c) 苏莱曼尼算死于暗杀吗?
Does the killing of Soleimani count as assassination?

(d) 日本是如何对待黑人与穆斯林的
How Japan treats the Black and Muslins.

(e) 日本海底核试验败露!地震后的奇异景象已证实?安倍这次该如何收场?
Japan's underwater nuclear test revealed! The bizarre scene after the earthquake has been confirmed? How should Prime Minister Abe react this time?

9. Others (OTH)

(a) 她不是一个人在战斗！
She is not fighting alone.

(b) 师恩·难忘
Teacher's kindness, unforgettable

(c) 国人为何热衷于炫耀?
Why are Chinese people keen on showing off?

(d) 最后，只有东北人没有被导游忽悠到
In the end, only the Northeast people were not fooled by the tour guide

(e) 马基雅维里——恶魔还是巨人?
Machiavelli — Demon or Giant?

Appendix E Supervised Text Analysis

In the supervised learning process, I only include the titles and authors of the articles in the text corpus because of two limitations of the dataset. First, as mentioned above, the URLs of the full articles did not work and therefore the data collection process of the full articles cannot be automated. Second, the formats of WeChat articles are inconsistent. Some include pure text whereas others have images and videos. Including content in the text analysis will lead to an unacceptable reduction rate of observations.

To predict the remaining 13,000 unlabeled data, I trained a multinomial logistic regression model with ridge estimator using the 2500 labeled data. Penalized regression models were used because the number of predictors (text) is much larger than the number of observations. Since I do not wish to drop predictors in the regularization process, L2 ("ridge") penalty is more preferable than L1 ("LASSO") penalty. The training model is specified as:

$$y_{ij} = \alpha_j + \mathbf{DFM}_i \beta_{\mathbf{j}} + \epsilon_{ij}$$

where y_{ij} is a binary variable that takes 1 if observation i belongs to topic category j and 0 otherwise. **DFM** is the document-feature matrix of the labeled data. **X** is a matrix of additional predictors. $\beta_{\mathbf{j}}$ is the matrix of ridge estimators for category j. Once the best matrices of ridge estimators, $\hat{\beta}_{\mathbf{j}}$, was found, I matched the unlabeled text corpus with the DFM of the labeled data. I then used the matched matrix and the best matrix of ridge estimators, $\hat{\beta}_{\mathbf{j}}$, to predict the unlabeled data.

Before the text analysis, all punctuation and stop words are removed and the Chinese text is segmented into individual tokens. Then, the segmented text was converted into a document-feature matrix. Words that appear less than 4 times were removed from the document-feature matrix. This produces a DFM with 1181 dimensions.

I also train two sets of alternative models to check the robustness of the predictions. They are linear regression models with ridge estimators and binomial logistic regression models with ridge estimators. The procedures are mostly the same as the multinomial ridge models. However, each topic category is predicted independently. The topic category with the highest predicted value is chosen to be the predicted topic category.

Table E1 shows that predictions are generally consistent across the three models. All three models predict political content to account for around 33-40% of all censored articles. Confirming the theoretical expectation that non-political content accounts for the majority of all censored content.

Table E1: Predicted Proportion of Censored Articles by Topic Category

	Multinomial	Linear	Binomial		Multinomial	Linear	Binomial
Political	39.34%	33.79%	38.46%	COL	0.67%	1.13%	1.01%
				CRI	27.94%	21.59%	27.69%
				GOV	10.73%	11.07%	9.76%
Non-Political	60.67%	66.22%	61.54%	BET	14.43%	16.36%	15.15%
				ESX	20.32%	26.21%	20.66%
				ADS	7.40%	8.00%	7.98%
				LCT	12.28%	11.21%	12.27%
				FOR	3.38%	3.77%	4.08%
				OTH	2.86%	0.67%	1.40%

Note: ADS: Advertisement. BET: Business & Economy. COL: Collective Action. CRI: Government Criticism. ESX: Entertainment & Sexuality. FOR: Foreign News. GOV: Government (non-criticism). LCT: Local Events, Culture, Traditions. OTH: Others.

Table E2: Top Keywords of Each Topic Category

Category	Keywords
Collective Action	暴徒 Violent Protesters 维权 Rightful Resistance 香港 Hong Kong 工会 Labor Union
Government Criticism	落马 Removed from Office 违规 Violating Laws 主任 Official 任志强 Ren Zhiqiang 纪委 Commission for Discipline Inspection
Government (Non-Criticism)	习近平 Xi Jinping 市长 Mayor 共青团中央 CCP Youth League 中方 Chinese side 改革开放 Reform and Openning Up
Business & Economy	谷歌 Google 马云 Jack Ma 股市 Stock Market 研发 Innovation 上市 IPO 京东 JD.com 资金 Capital 网易 NetEase Inc
Entertainment & Sexuality	老婆 Wife 单身 Single 吴秀波 Wu Xiubo 爱情 Love 娱乐圈 Entertainment Circle
Advertisement	优质 Good Quality 轻松 Easy 睡眠 Sleeping 肠胃 Stomach 蜂蜜 Honey
Local Events & Traditions & Cultures	景区 Tourist Attractions 道德 Morality 公交 Public Transportation 中医 Chinese Medicine
Foreign News	朝鲜 North Korea 全世界 The World 日本 Japan 特朗普 Donald Trump 土耳其 Turkey

Note: Keywords are extracted from linear ridge models and binomial ridge models. Keywords are among the top predictors of each topic category.

Appendix F Validation of Text Analysis

To validate the multinomial ridge model prediction, I first conduct an in-sample validation. Table F1 reports the Confusion Matrix between human coder 1 and multinomial model prediction using the 2,500 articles in the training set. The accuracy rate of the multinomial model is 81.64% and the macro F1 is 0.80. The accuracy rate of the model in identifying political content in general is 90.04% and the macro F1 is 0.89 (see table F2). The Cohen's κ between coder 1 and model prediction is 0.79.

According to table F1, government criticism (CRI) is over-estimated by the multinomial model, whereas most other categories are under-estimated. This also indicated by the high recall rate of government criticism (CRI) and high precision rate of other categories. Table F2 confirms that political content is over-estimated by the multinomial model. Hence, the bias is against my theoretical prediction, and the proportion of political content reported in the paper is likely to be the upper-bound.

I further conduct three-fold cross-validation using the multinomial model predictions. I randomly divide the 15,872 articles into three groups. Each time, I use two groups as the training set to predict the remaining group. The training and prediction process is the same as the main analysis (see section 5).

Table F3 and F4 report the Confusion Matrixs of cross-validation results. In general, the cross-validation results are consistent with the in-sample validation results. The accuracy rate of the cross-validation is 76.47% for specific topic categories and 85.87% for general category. The macro F1 for specific topic categories is 0.80 and the macro F1 for general topic categories is 0.89. Government Criticism (CRI) is also over-estimated by the cross-validation, which leads to the over-estimation of political content in general. Despite such bias, non-political content still accounts for the majority of the censored articles. The Cohen's κ between coder 1 and model prediction is 0.71.

Table F1: Confusion Matrix: In-Sample Validation of the Multinomial Ridge Model Prediction

		Multinomial Model Prediction									
		ADS	BET	COL	CRI	ESX	FOR	GOV	LCT	OTH	
Human Coder 1	ADS	188	8	0	18	24	1	5	5	1	
	BET	3	313	0	22	3	0	1	1	2	
	COL	0	3	40	9	9	1	1	5	1	
	CRI	2	12	0	382	11	3	7	5	0	
	ESX	1	3	0	22	408	1	1	3	1	
	FOR	2	2	0	34	8	112	3	5	1	
	GOV	0	11	1	41	5	1	243	8	1	
	LCT	3	9	0	16	13	1	5	260	4	
	OTH	6	9	0	40	22	0	4	9	95	
		ADS	BET	COL	CRI	ESX	FOR	GOV	LCT	OTH	Macro
	Precision	0.92	0.85	0.98	0.65	0.81	0.93	0.90	0.86	0.90	0.87
	Recall	0.75	0.91	0.58	0.91	0.93	0.67	0.78	0.84	0.51	0.76
	F1	0.83	0.88	0.73	0.76	0.87	0.78	0.84	0.85	0.65	0.80

Note: ADS: Advertisement. BET: Business & Economy. COL: Collective Action. CRI: Government Criticism. ESX: Entertainment & Sexuality. FOR: Foreign News. GOV: Government (non-criticism). LCT: Local Events, Culture, Traditions. OTH: Others.

Table F2: Confusion Matrix: In-Sample Validation (General Category)

		Multinomial Model Prediction		
		Political	Non-Political	Macro
Human Coder 1	Political	724	78	
	Non-Political	171	1527	
	Precision	0.81	0.95	0.88
	Recall	0.90	0.90	0.90
	F1	0.85	0.92	0.89

Note: Political content includes collective action (COL), Government Criticism (CRI), and Government Non-Criticism (GOV). Other categories are non-political content.

Table F3: Confusion Matrix: Three-Fold Cross-Validation (Specific Category)

		Cross-Validation Prediction									
		ADS	BET	COL	CRI	ESX	FOR	GOV	LCT	OTH	
Multi-nomial Model	ADS	849	43	0	115	98	6	16	37	10	
	BET	30	1739	0	276	74	12	67	79	13	
	COL	0	5	43	33	7	0	6	12	0	
	CRI	28	145	1	3815	184	24	126	98	13	
	ESX	43	58	1	352	2632	8	38	85	8	
	FOR	8	36	0	128	37	274	23	22	9	
	GOV	15	74	1	247	60	22	1182	97	5	
	LCT	25	71	0	231	108	9	73	1418	14	
	OTH	12	24	0	111	63	10	22	26	186	
		ADS	BET	COL	CRI	ESX	FOR	GOV	LCT	OTH	Macro
	Precision	0.84	0.79	0.93	0.72	0.81	0.75	0.76	0.76	0.72	0.79
	Recall	0.72	0.76	0.41	0.86	0.82	0.51	0.69	0.73	0.41	0.66
	F1	0.78	0.76	0.57	0.78	0.81	0.61	0.73	0.74	0.52	0.70

Note: Rows are multinomial model predictions from supervised text analysis. Columns are cross-validation predictions from the results. ADS: Advertisement. BET: Business & Economy. COL: Collective Action. CRI: Government Criticism. ESX: Entertainment & Sexuality. FOR: Foreign News. GOV: Government (non-criticism). LCT: Local Events, Culture, Traditions. OTH: Others.

Table F4: Confusion Matrix: Three-Fold Cross-Validation (General Category)

		Cross-Validation Prediction		
		Political	Non-Political	Macro
Multinomial Prediction	Political	5454	789	
	Non-Political	1453	8176	
	Precision	0.79	0.91	0.85
	Recall	0.87	0.85	0.86
	F1	0.83	0.88	0.86

Note: Political content includes collective action (COL), Government Criticism (CRI), and Government Non-Criticism (GOV). Other categories are non-political content.

Appendix G Other Sampling Concerns of the Censorship Data

G.1 Non-representative Public Accounts

The first concern is that the selected public accounts are not representative of all public accounts, which creates bias. Using the name and introduction of the public accounts, I manually identified five different types of accounts: political, business, entertainment, foreign, and local accounts. Accounts that do not fall into the definition of any type will be labeled as other accounts. The coding rubrics generally resembles the rubrics for individual articles (see section 5.2). Table G1 shows the numbers and percentages of each account type, as well as the numbers and percentages of articles published by each account type. By both measures, political accounts is by far the largest account type. Hence, even if the current sample is not representative, it likely records higher proportion of political censorship than the true population and any bias is against my theoretical expectation.

Table G1: Summary Statistics of Account Types

Account Type	Political	Business	Foreign	Entertain	Local	Others	Total
# of Account	779	176	67	137	309	438	1906
% of Account	0.409	0.093	0.035	0.162	0.162	0.230	1
# of Articles	8402	1299	287	1448	2207	2229	15872
% of Articles	0.529	0.082	0.018	0.091	0.139	0.140	1

Table G2 reports the predicted proportion of the 8402 censored articles published by political accounts. Non-political content still consists of a substantial proportion, just slightly lower than 50%. Therefore, even people who only subscribe to political accounts still have a fair chance of encountering censorship of apolitical content.

Table G2: Predicted Proportion of Censored Articles Published by Political Accounts

General Category	Percentage	Specific Topic Category	Percentage
Political	50.04%	Collective Action	1.13%
		Government Criticism	37.56%
		Government (Non-Criticism)	11.71%
Non-Political	49.59%	Business & Economy	12.90%
		Entertainment & Sexuality	13.35%
		Advertisement	5.43%
		Local Events, Traditions, Cultures	9.47%
		Foreign News	4.99%
		Others	3.45%

G.2 Government Censorship Intention

Another concern of the censorship data is that some of the censored articles might not reflect government intention. The WeChatScope dataset records the official censored reason of each article. Around 4,637 of the 15,872 articles are recorded as "Violating Internet Laws" or "Account Blocked", whereas the remaining 11,235 articles are recorded as "Deleted by the Author". However, the official censored reason is not a reliable measure of why the articles were censored, and distinguishing between government censorship and self-censorship is not simple. The authors of the WeChat articles might be forced to delete their articles by their boss, local party leaders, or other government officials. WeChat can also tell the authors to take down their articles. In some cases, similar articles on the same incident were censored in a relatively short period, but some of the articles were recorded as "Violating Internet Laws", whereas others were recorded as "Deleted by the Author". This further shows government censorship might happen even when the official censored reason is "Deleted by the Author".

Moreover, because my theory focuses on the impact of the government's censorship strategy on citizens' reactions, it does not matter whether such a goal is achieved by directly censoring the articles, or forcing the authors to delete their articles. From the citizens' perspective, the signals they receive are the same, regardless of the official censored reason.

Nevertheless, I subset my dataset to those articles that are recorded as "Violating Internet Laws" or "Account Blocked". This is the most conservative estimation strategy and all articles included are without any doubt reflecting government censorship intention. As shown in the table G3, even using this extremely conservative estimation strategy, non-political content still consists of 46.05% of all censored articles. It is still plausible to claim that non-political censorship happens on a substantial scale and censorship normalization does happen in China.

Table G3: Predicted Proportion of Censored Articles by Topic Category

General Category	Percentage	Specific Topic Category	Percentage
Political	53.93%	Collective Action	1.23%
		Government Criticism	43.30%
		Government (Non-Criticism)	9.40%
Non-Political	46.05%	Business & Economy	13.24%
		Entertainment & Sexuality	13.11%
		Advertisement	4.10%
		Local Events, Traditions, Cultures	8.34%
		Foreign News	4.59%
		Others	2.67%

Note: The total number of articles is 4,637. It only includes articles that are recorded as "Violating Internet Laws" or "Account Blocked".