

Appendices

Appendix A. Variables, measurement and inter-rater reliability

Variable	Coding	Range	Mean (SD)	Brennan & Prediger's κ (% agreement) ¹
<i>Dependent variables</i>				
Source type	Source in which protest was recorded (one protest event can be recorded in multiple sources)			
Social media	Chinese social media	0/1	0.93 (0.25)	
Dissident media	Dissident websites	0/1	0.90 (0.29)	
International media	Chinese and English international news media	0/1	0.03 (0.17)	
Domestic media	Domestic news media	0/1	0.01 (0.11)	
Protests per day (social media)	The number of protest events that occurred in a city on a given day (city-day)	0 – 23	1.09 (1.30)	
Protests per day (dissident media)	The number of protest events that occurred in a city on a given day (city-day)	0 – 4	0.11 (0.34)	
Protests per day (news media)	The number of protest events that occurred in a city on a given day (city-day)	0 – 2	0.04 (0.21)	

¹ For media content-based variables only. After sufficient interrater reliability was established during coder training, it was re-tested on five occasion and altogether 139 cases over the course of the coding process. The reported reliability measures refer to these 139 cases, 119 of which were coded by coders A and B and 20 by coders A and C.

Protest participants per day (social media)	The number of estimated protest participants in a city on a given day (city-day)	0 – 26500	118.17 (630.42)	
Protest participants per day (dissident media)	The number of estimated protest participants in a city on a given day (city-day)	0 – 26500	55.44 (590.63)	
Protest participants per day (news media)	The number of estimated protest participants in a city on a given day (city-day)	0 – 26500	39.16 (585.41)	
Average protest participants per day (social media)	The number of estimated protest participants in a city on a given day (city-day)	0 – 26500	75.93 (572.41)	
Average protest participants per day (dissident media)	The number of estimated protest participants in a city on a given day (city-day)	0 – 26500	51.08 (578.90)	
Average protest participants per day (news media)	The number of estimated protest participants in a city on a given day (city-day)	0 – 26500	38.00 (576.70)	
<i>Independent variables</i>				
Event	The event qualified as a protest event according to the definition. It occurred within the jurisdiction of Guangzhou, Chongqing, or Shanghai.	0/1	0.92 (0.27) ²	0.92 (96%)
Large protests	Protest with 500 participants or more	0/1	0.04 (0.19)	0.98 (99%)
Protester violence	Protesters use violence against people	0/1	0.03 (0.18)	0.93 (96%)
State coercion	The police used, or threaten the use of, coercive methods.	0/1	0.25 (0.43)	0.90 (95%)

² All other summary statistics refer to the data set with cases where event = 0 have been dropped.

Discontent with the state	This variable was combined from variables on discontent over political procedure, public policy, miscarriage of justice, and official misconduct. ³	0/1	0.14 (0.34)	0.89 (95%)
Discontent with the regime	Protests voicing demands for political (elections) and civic rights (freedom of expression, assembly, association, due process), human rights, official asset disclosure, election of labor union representatives, support for human rights activists and other dissidents.	0/1	0.02 (0.14)	0.96 (98%)
Series	The event was preceded by another separate protest event in the same location, with the same group of participants for the same issue.	0/1	0.12 (0.32)	
Days	Number of days an event lasted.	1-39	1.15 (1.18)	
Protest participants	Average of maximum and minimum protest participant estimates	6-26500	112.55 (580.72)	
City	City in which protest took place	0/1		
Shanghai		0/1	0.31 (0.46)	
Guangzhou		0/1	0.29 (0.46)	
Chongqing		0/1	0.39 (0.49)	

³ Coding instructions were: “Legitimate political procedure” (Allegations of misconduct or violation of procedures in election of residential committee/homeowners’ committee, other elections, public consultations/hearings, environmental impact assessment, project approval, compensation determination, information release, etc.); “Public policy, administrative act, regulations, fees” (Discontent with a specific government policy, administrative act, law, regulation, or fee.); “Miscarriage of justice, public security” (Discontent with the conduct/outcome of, or failure to file, a lawsuit, police/urban administration brutality or misconduct, illegitimate arrests or police investigations, conflicts over evidence (such as dead bodies), failure to combat crime, alleged police-crime collaboration, etc.); “Official corruption and misconduct” (Allegations of officials having used public office for private gain, embezzled public funds, had improper sexual relationships, acted tyrannically etc. There must be evidence that protesters made accusations of corruption.).

Public holidays	National public holidays according to central government announcements. ⁴	0/1	0.02 (0.14)
Rainfall	Daily amount of rainfall in mm.	0-165.1	9.01 (16.48)
No. workers' protests (social media)	The number of workers' protest events that occurred in a city on a given day (city-day).	0 – 21	0.40 (0.92)
No. workers' protests (dissident media)	The number of workers' protest events that occurred in a city on a given day (city-day).	0 – 1	0.02 (0.15)
No. workers' protests (news media)	The number of workers' protest events that occurred in a city on a given day (city-day).	0 – 2	0.02 (0.13)
Weekday	Week day in which protest occurred	0 – 6	3.00 (2.00)
Month	Calendar month in which protest occurred	1 – 12	5.92 (3.46)
Month counter	Variable measuring year and month in which a protest occurred	1-29	15.02 (8.36)

⁴ See, http://www.gov.cn/zhengce/content/2015-12/10/content_10394.htm; http://www.gov.cn/zwggk/2013-12/11/content_2546204.htm; http://www.gov.cn/zhengce/content/2014-12/16/content_9302.htm, last accessed March 13, 2020.

Appendix B. Search keywords

The keywords used for Chinese sources were generated by first sending a list of 30 terms to 5 experts on contentious politics in China (4 native speakers and 1 non-native speaker) to ask for comments. A subsequently amended list of 58 terms was then manually tested on 1 dissident medium, 7 Mainland Chinese newspapers and 7 Chinese newspapers from Hong Kong and Taiwan for coverage and efficiency. The following 25 terms were thereby extracted:

群体性事件 quntixing shijian (mass incident), 群体事件 qunti shijian (mass incident), 示威 shiwei (demonstration), 游行 youxing (parade), 骚乱 saoluan (riot), 集会 jihui (assembly), 上街 shangjie (take to the streets), 静坐 jingzuo (sit-in), 集体上访 jiti shangfang (collective petition), 请愿 qingyuan (petition), 闹事 naoshi (trouble-making), 扰乱社会秩序 raoluan shehui zhixu (disrupt social order), 扰乱公共秩序 raoluan gonggong zhixu (disrupt public order), 罢课 bake (student/teacher strike), 罢市 bashi (shopkeeper's strike), 讨说法 讨个说法 tao (ge) shuofa (demand an explanation), 停工 ting gong (work stoppage), 罢工 bagong (strike), 讨薪 taoxin (bargain salaries), 暴力抗法 baoli kangfa (violent resistance against law enforcement), 喊口号 han kouhao (shout slogans), 横幅 hengfu (banner), 警民冲突 jingmin chongtu (police–people conflict), 催泪弹 cui lei dan (tear gas canisters)

The keywords used for LexisNexis followed Weidman and Rød:⁵ “protest”, “demonstration”, “rally”, “campaign”, “riot” and “picket”. We added the terms “strike” and “unrest” to these terms.

⁵ Nils B. Weidmann and Espen Geelmuyden Rød, *The Internet and Political Protest in Autocracies* (Oxford: Oxford University Press, 2019).