Elite Competition, Local Extraction, and Social Unrest: Understanding Mass Protest in Authoritarian Regimes Online Appendix

A The Tongan Protest in Jiangsu Province

To understand why collective resistance occurs in China, I describe how land expropriation by Suzhou City in Jiangsu, motivated by local competition, incited massive resistance to illustrate the chain of influence and provide some intuition for the base of the explanations described in the previous section. The following news extracted from the ICEWS database reports a land-related protest in Jiangsu province on July 16, 2010:

> "The protest started on July 14 when thousands in Jiangsu province's Suzhou city gathered at the government headquarters of Tongan town, Caijing Magazine reported on its website. Thousands of people in eastern China protested for at least five days against local authorities which they accuse of confiscating their land illegally and withholding land compensation. Approximately 1,000 villagers from Tongan township have been gathered outside the government building since 14 July and these residents clashed with riot police on Friday evening when some protesters were injured and others were taken away. ... The next day, thousands assembled on a highway and blocked traffic, but were later dispersed by police. ... Protesters returned on Sunday, their ranks swelled to 10,000. They also dispersed, the report said."

Land expropriation in Tongan was a part of a larger integrated effort authorized by the Suzhou City government and executed by local towns and villages to convert local farmlands for the purpose of expanding industrial park constructions. Since 2003, the Suzhou City government has confiscated large amounts of farmland to attract outside investment and stimulate industrial development. These local residential and agricultural land conversion efforts attracted 70 billion USD in investment capital to the region.[1] To accelerate the speed and the scale of land conversion in local towns and counties, and thus increase the land supply for industrial constructions, in 2003 the Suzhou City government began a merger of the Gaoxin and Huqiu districts.[2] Seeking further economic benefit, Suzhou city initiated

[1] `http://www.gmw.cn/01gmrb/2007-09/23/content_674898.htm`

[2] Districts in China are equivalent to the level of county administrative units.

a new round of extensive local land confiscation in 2010 to convert approximately 12.5 square kilometers and commanded subordinate governments to integrate their efforts in pushing land conversion processes following the principle of *strong government* in the South Jiangsu Economic Model. The initiative put pressure on local townships to accelerate the conversion of vast amounts of rural farmland for industrial usage, and the corresponding relocation policy affected more than 9000 villagers across multiple localities such as Tongan, Xushuguan, and Dongzhu. Large-scale confiscation and unsatisfactory settlement triggered several noticeable waves of collective resistance in Tongan township and neighboring areas, including the towns of Xushuguan and Dongzhu. Land-losing residents filed petitions to Suzhou City government, but responses never reached the residents. The enraged protesters claimed that they were forced to be relocated and the local government did not consult them before undertaking massive land confiscations. The compensation that was promised was either delayed, underpaid, or never given to residents. According to Tongan residents, a legal document issued in 2008 revealed that every relocated unit should have received compensation of 400,000 RMB, but residents said that none of this was received. They filed their complaint to the township government, but the township officials responded that these land transfers were legal actions that were approved by the upper-level government for technology development zones and the current administration was unwilling to repay the debt left by its predecessor.[3] Residents were angry in response and participated in ensuing protests and riots.

The Tongan protest was not an isolated incident but exemplified severe land injustice resulting from local competition among cities in Jiangsu province. To embellish their performance and signal political achievement, local governments not only increased their land provision but also competed over how much they could lower land prices to attract investors. Luo and Lin (2003) discovered in field research that the heated competition over development in south Jiangsu province resulted in the comprehensive underpricing of land leased in Xuzhou City, Changshu City, and Wuxi City. Once investment had taken off, the average value of land in Jiangsu was estimated at around 200,000 RMB per acre, but the actual leasing price was only in the range of 80-120 thousand RMB per acre (Shen, 2006). The competitive land price depreciation and the corresponding reduction in lease earnings for local governments often resulted in undercutting relocation subsidies for land-losing residents and incited resistance. To prevent aggravated opportunistic behaviors and regulate the distorted land leasing market, the State Council even enforced a minimum leasing price for cities around the Yangtze River Delta Region.[4] But this minimum leasing price did not resolve local land hunger, nor did it dissuade officials from pursuing the fiscal and political benefit incurred from foreign investment because lands were seen as a strong attraction for capital investment. Unregulated land conversion was also aggravated by defective land confiscation laws and unsupervised converting procedures, which allowed local agents to find loopholes in the system to underpay compensation during expropriation. Taken together, politically motivated land takings coupled with defective land-transfer supervision resulted in severe expropriation problems in China and threatened social stability.

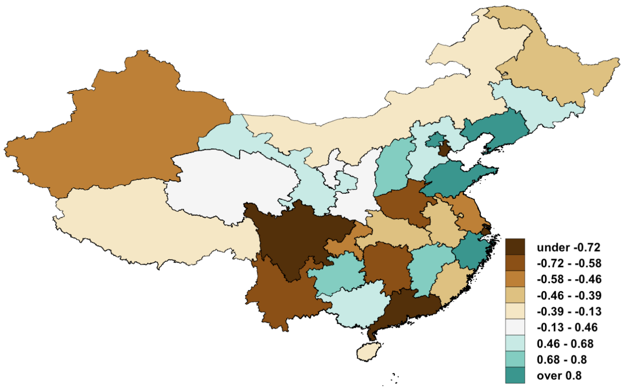
[3] `http://news.163.com/10/0819/04/6EE3R30A00014AED.html`

[4] `http://news.xinhuanet.com/fortune/2008-09/25/content_10107179.htm`

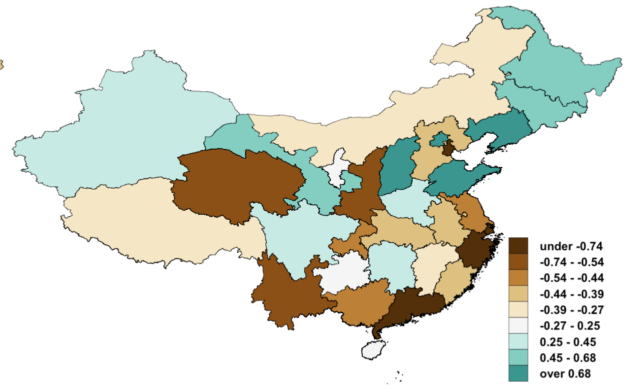
B Residual maps, 2001-2012

Figure A.1: Residual Maps, 2001-2012

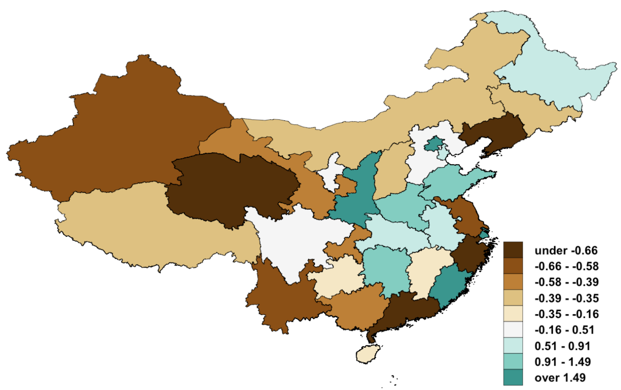
(a) 2001



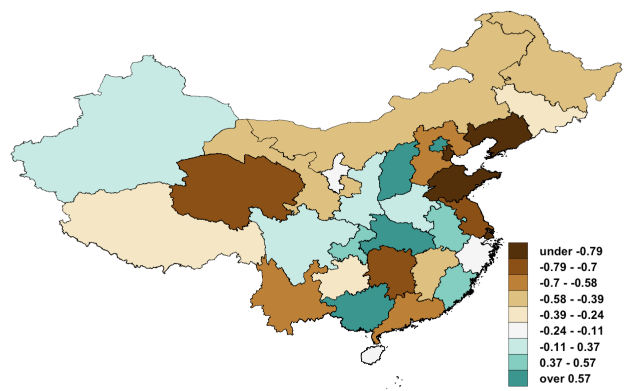
(b) 2002



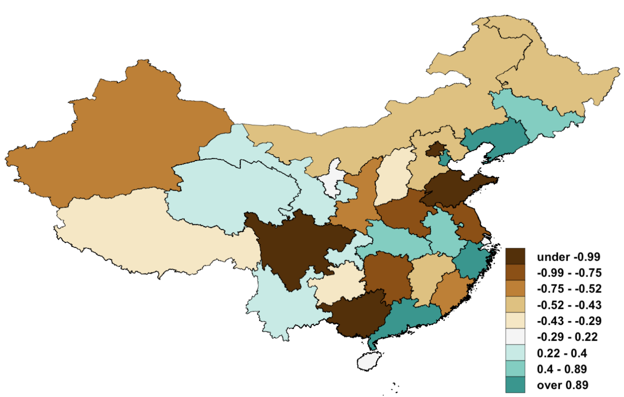
(c) 2003



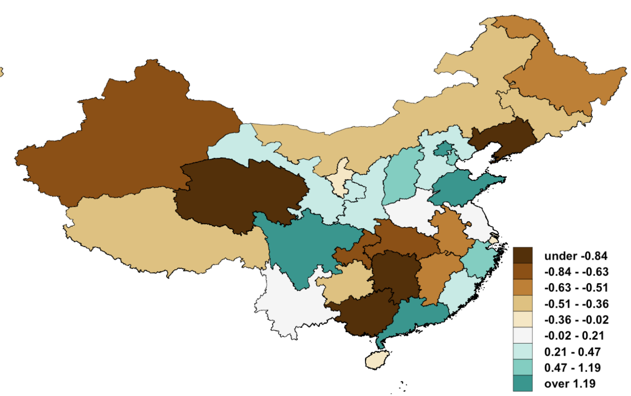
(d) 2004

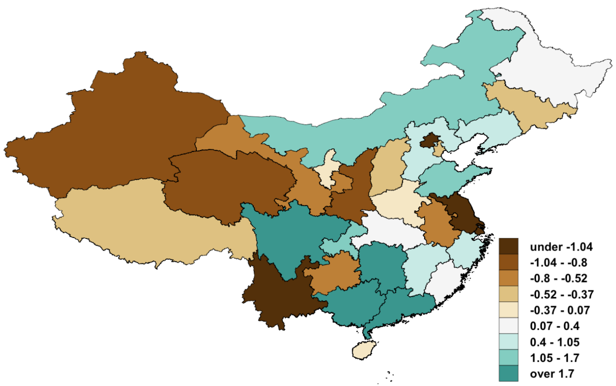


(e) 2005

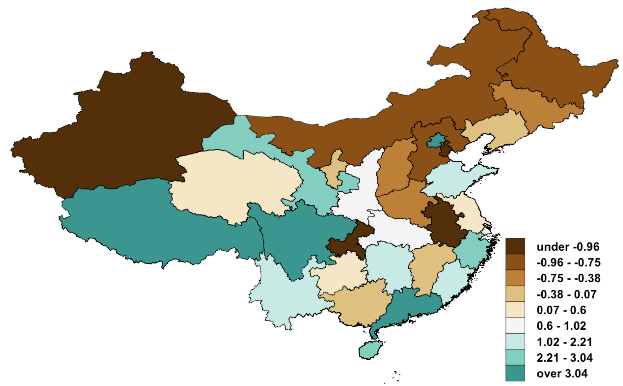


(f) 2006

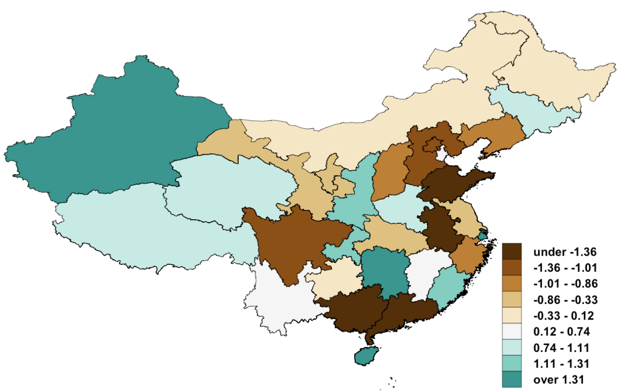




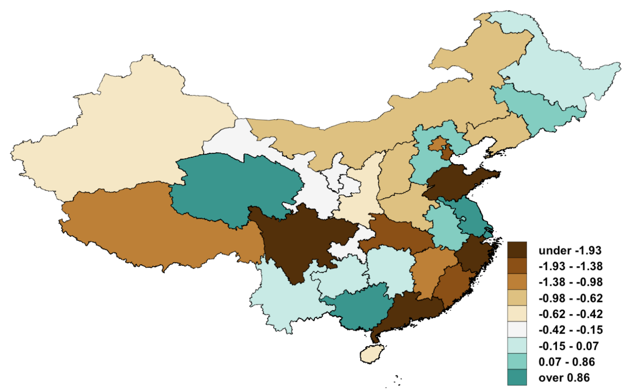
(g) 2007



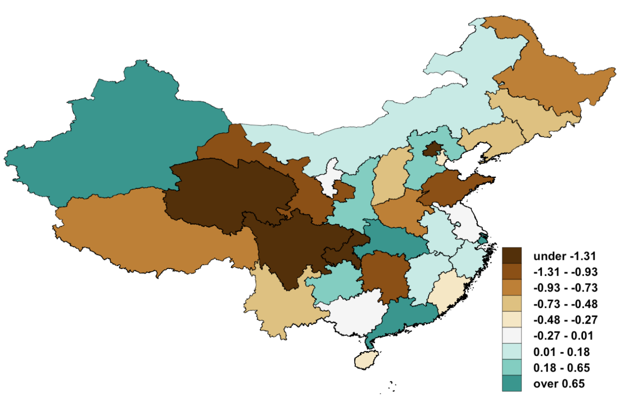
(h) 2008



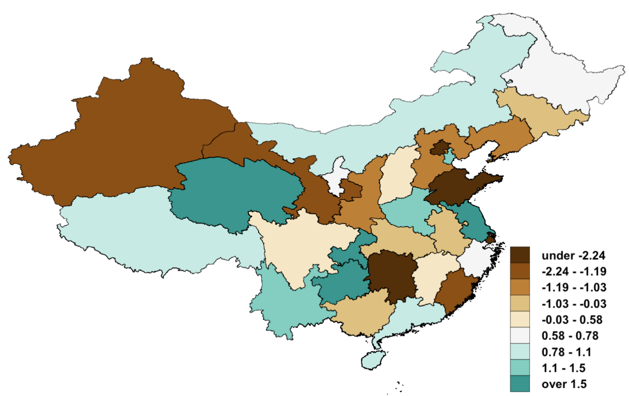
(i) 2009



(j) 2010



(k) 2011



(l) 2012

C Protest Events Recorded by ICEWS

Figure A.1: Total Sum of Protests Across Provinces in China, 2001-2014

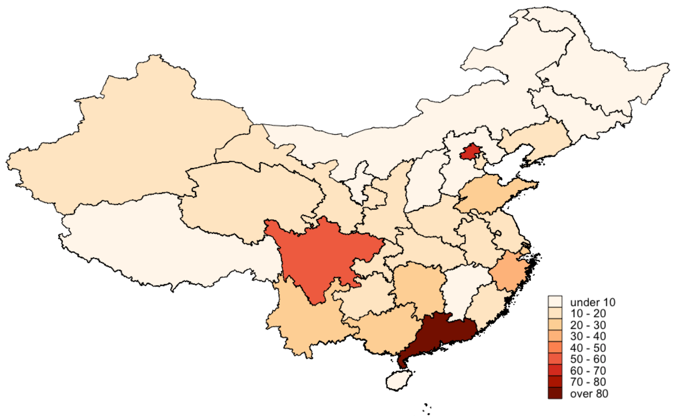
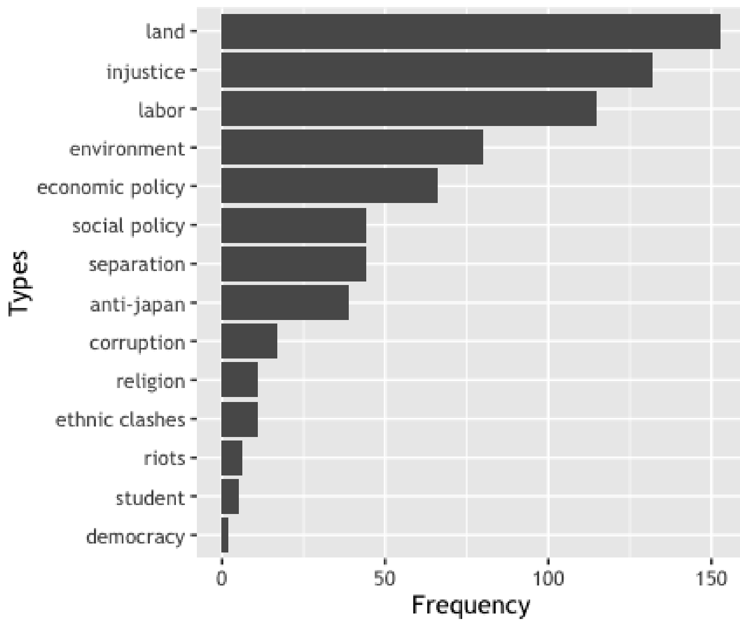


Table A.1: Dataset Comparison

	province	ICEWS	CASS	Tong
1	anhui	10	21	18
2	beijing	75	35	13
3	chongqing	15	15	21
4	fujian	13	22	6
5	gansu	11	18	10
6	guangdong	138	267	135
7	guangxi	23	29	19
8	guizhou	8	17	5
9	hainan	8	23	10
10	hebei	7	19	10
11	heilongjiang	7	11	10
12	henan	19	39	23
13	hubei	15	19	26
14	hunan	21	29	19
15	inner mongolia	5	6	5
16	jiangsu	20	43	27
17	jiangxi	9	27	14
18	jilin	7	9	6
19	liaoning	12	7	11
20	ningxia	2	1	1
21	qinghai	20	7	3
22	shaanxi	17	22	19
23	shandong	19	19	25
24	shanghai	30	22	8
25	shanxi	5	14	10
26	sichuan	52	54	41
27	tianjin	13	5	4
28	tibet	9	0	3
29	xinjiang	13	5	5
30	yunnan	19	32	17
31	zhejiang	25	35	21

D Event Issue Types

Figure A.2: Event Types of Protest in China, 2001-2014



Data source: ICEWS protest event data

Figure A.2 shows protest frequencies in different event types. I coded these event types following protest categories organized by (Tong and Lei, 2013). Their categorization is based on the nature of the conflict and protesters' claims. The land protest category contains issues related to land expropriation by the governments. Injustice protest contains mainly about brutal and unlawful treatment by police. Labor protest contains events targeting state-own enterprises (SOEs) or private companies by workers. Environmental protest refers to disruptions targeting pollution. The economic-policy category contains protests aiming at unfair taxation-and-fee policies, while social policy contains mass events targeting things such as language policies or the one-child policy. Separation protests are those targeting at separatism like Tibetan and Xinjiang Uygur resistances. Anti-Japan protests refer to events targeting the island disputes and associated nationalism protests. It should be noted that ICEWS treats anti-Japan protests that escalated to the nationwide level as an integrated event due to the temporal proximity and the nature of the goal. For example, two main waves of protests in mid-August and mid-September 2012 were coded as two mass events. These widespread nationalism events are rare in China (due to the encouragement by the state); most protests are very locally concentrated. Religion protests are mostly on Falun Gong persecution. Ethnic clashes refer to the conflict between different ethnicities, while mass riots are those pertaining to fights between gangs or citizens. Democracy refers to protests claiming human rights improvement and open democracy, which are mostly located in Beijing.

E Mass Protest Summary Report by CASS

Figure A.3: Mass Protest Summary of Frequency by CASS, 2000-2013

表 2　2000 年 1 月 1 日至 2013 年 9 月 30 日中国境内群体性事件的地域分布

单位：件

区域	项目 / 省份	百人至千人群体性事件数量	千人至万人群体性事件数量	万人以上群体性事件数量	总数量
华北	北京	32	3	0	35
	天津	1	4	0	5
	内蒙古	5	1	0	6
	河北	15	4	0	19
	山西	13	1	0	14
	总计	66	13	0	79
华东	上海	13	9	0	22
	山东	13	6	0	19
	江苏	25	16	2[①]	43
	浙江	22	12	1[②]	35
	安徽	14	7	0	21
	江西	19	8	0	27
	福建	12	10	0	22
	总计	118	68	3	189
华中	河南	24	15	0	39
	湖北	11	7	1[③]	19
	湖南	17	12	0	29
	总计	52	34	1	87
华南	广东	189	76	2[④]	267
	广西	19	10	0	29
	海南	15	8	0	23
	总计	223	94	2	319
东北	黑龙江	6	5	0	11
	吉林	7	1	1[⑤]	9
	辽宁	3	3	1[⑥]	7
	总计	16	9	2	27
西北	甘肃	17	1	0	18
	宁夏	0	1	0	1
	青海	5	2	0	7
	陕西	19	3	0	22
	新疆	5	0	0	5
	总计	46	7	0	53
西南	四川	25[⑦]	28	1[⑧]	54
	贵州	8	8	1[⑨]	17
	云南	27	5	0	32
	重庆	10	5	0	15
	总计	70	46	2	118

F Simulating the Effect of Potential Media Bias on Protest Reports

Regarding potential reporting biases, I control three additional variables in the models: Internet penetration rate, number of local (printed) newspapers, and local distance to the country capital, Beijing. The Internet penetration rate in provinces aims to capture the probability of protest events being reported to the web and thereby recorded by external media outlets in Beijing, Hong Kong, or other regional and international media, while the number of printed newspapers estimates event reporting capacity of media in general.[5] The distance to the capital measures the level of attention likely to be given to local events in the mainstream media outlets. Table A.2 shows that the result is still robust after considering these several potential sources of reporting bias.

Table A.2: Results by Controlling Potential Reporting Bias (Land Protests)

	Model 1	Model 2	Model 3	Model 4
Intercept	-1.522^{***}	-1.449^{***}	-1.421^{***}	-1.442^{***}
	(0.160)	(0.167)	(0.156)	(0.158)
Leaders	2.238^{***}	1.879^{***}	1.997^{***}	1.464^{**}
	(0.595)	(0.580)	(0.611)	(0.609)
Leaders^2	-1.808^{***}	-1.498^{***}	-1.547^{***}	-1.078^{**}
	(0.527)	(0.513)	(0.532)	(0.537)
Ln area	0.260	−0.084	0.021	−0.008
	(0.215)	(0.212)	(0.205)	(0.206)
Ln population lag	−0.647	−0.205	−0.053	−0.070
	(0.461)	(0.433)	(0.446)	(0.447)
Ln GDP lag	0.195	0.129	−0.234	−0.282
	(0.464)	(0.512)	(0.480)	(0.475)
Urbanization lag	0.159	0.135	−0.105	−0.035
	(0.263)	(0.276)	(0.357)	(0.309)
Security expense lag	0.267	0.296	0.411^{**}	0.445^{***}
	(0.170)	(0.185)	(0.173)	(0.170)
Protest time lag	0.035	0.063	0.069	0.078
	(0.076)	(0.082)	(0.081)	(0.080)
Protest spatial lag	0.185	0.178	0.287^{**}	0.310^{**}
	(0.115)	(0.121)	(0.123)	(0.123)
Distance to HK	-0.567^{***}			
	(0.190)			
Distance to Beijing		0.374^{**}		
		(0.181)		
Internet penetration rate			0.306	
			(0.200)	
News agencies				0.312^{**}
				(0.135)
AIC	460.692	464.936	467.308	464.182
BIC	509.183	513.427	515.799	512.673
Log Likelihood	-217.346	-219.468	-220.654	-219.091
Num. obs.	308	308	308	308
Num. groups: province	28	28	28	28

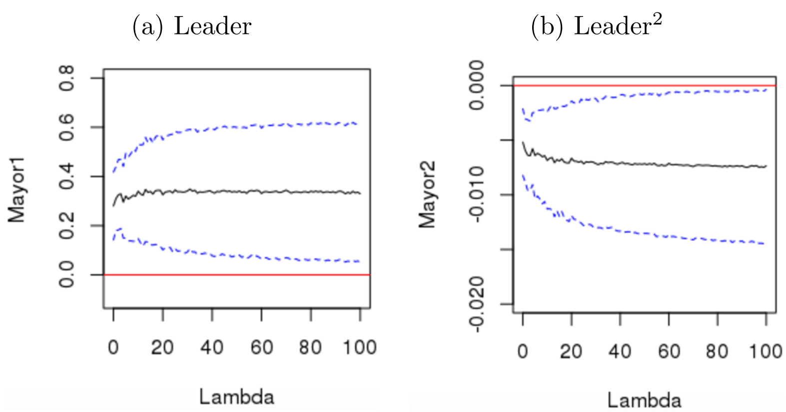
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

In addition to adding control variables, a simulation-based sensitivity analysis measures the potential bias from a different methodological perspective. Although media bias in news

[5] I do not consider how popular these printed newspapers are (i.e., their circulation) because this information is hard to obtain.

report data has been widely discussed, there is no standard way of testing how sensitive a result is, and it is particularly difficult when the true record of events is not available (Weidmann, 2015). Therefore, in addition to comparing the data with other data sources and controlling potential sources of bias, I simulated more events to potentially under-reported areas to see how sensitive our result is. The method introduced by Gallop and Weschle (2019) suggests simulating random draws from Poisson distribution and re-estimating parameters in event count models.[6] Based on this rule, I simulated up to 100 times with the data in the top 1, 5, 10 underreporting regions. Figure A.4 shows our estimates in top 10 underreporting provinces. The first term and the quadratic term of the competition variable still hold across the number of simulations, suggesting that our result is less likely to be affected even if the bias exists.

Figure A.4: Results with Simulated Counts of Protest Events



[6] More mathematical details can be found in their paper.

G Appendix: Average Number of City Leaders in Provinces

Table A.3: Average Number of City Leaders

	Province	Average
1	anhui	17
2	beijing	18
3	chongqing	40
4	fujian	9
5	gansu	14
6	guangdong	21
7	guangxi	14
8	guizhou	9
9	hainan	2
10	hebei	11
11	heilongjiang	13
12	henan	17
13	hubei	13
14	hunan	14
15	inner mongolia	12
16	jiangsu	13
17	jiangxi	11
18	jilin	9
19	liaoning	14
20	ningxia	5
21	qinghai	8
22	shaanxi	10
23	shandong	17
24	shanghai	18
25	shanxi	11
26	sichuan	21
27	tianjin	17
28	tibet	7
29	xinjiang	14
30	yunnan	16
31	zhejiang	11

H Appendix: Result Tables with Increasing Numbers of Covariates

Table A.4: Random Effect Model (All Protest)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Intercept	0.213*	0.214*	0.213*	0.078	0.067	0.089	0.100	0.066
	(0.117)	(0.115)	(0.116)	(0.116)	(0.117)	(0.107)	(0.095)	(0.109)
Leaders	1.419***	1.402***	1.311***	0.972**	1.241***	1.310***	1.134***	1.416***
	(0.358)	(0.353)	(0.431)	(0.425)	(0.456)	(0.421)	(0.380)	(0.434)
Leaders^2	−1.096***	−1.099***	−1.024**	−0.672*	−0.900**	−0.970**	−0.841**	−1.051***
	(0.352)	(0.348)	(0.403)	(0.399)	(0.424)	(0.389)	(0.349)	(0.396)
Ln area		−0.105	−0.111	0.186	0.040	−0.021	−0.004	−0.063
		(0.115)	(0.117)	(0.119)	(0.148)	(0.139)	(0.126)	(0.145)
Ln population lag			0.053	−0.769***	−1.028***	−0.909***	−0.873***	−0.894***
			(0.145)	(0.171)	(0.233)	(0.229)	(0.216)	(0.259)
Ln GDP lag				1.073***	1.306***	0.969***	1.007***	0.935***
				(0.123)	(0.188)	(0.247)	(0.246)	(0.300)
Urbanization lag					−0.318*	−0.324*	−0.311*	−0.361*
					(0.191)	(0.181)	(0.171)	(0.205)
Security expense lag						0.212**	0.094	0.133
						(0.106)	(0.113)	(0.119)
Protest time lag							0.184***	0.100
							(0.068)	(0.070)
Protest spatial lag								0.141**
								(0.067)
AIC	1219.874	1221.054	1222.918	1152.090	1151.046	1149.002	1142.550	1070.106
BIC	1239.468	1244.567	1250.350	1183.441	1186.316	1188.191	1185.658	1116.089
Log Likelihood	-604.937	-604.527	-604.459	-568.045	-566.523	-564.501	-560.275	-523.053
Num. obs.	372	372	372	372	372	372	372	341
Num. groups: province	31	31	31	31	31	31	31	31

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table A.5: Random Effect Model (28 Provinces and Extraction-related Protest)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Intercept	-0.542^{***}	-0.548^{***}	-0.583^{***}	-0.624^{***}	-0.625^{***}	-0.549^{***}	-0.529^{***}	-0.562^{***}
	(0.134)	(0.134)	(0.138)	(0.134)	(0.133)	(0.109)	(0.092)	(0.102)
Leaders	1.606^{***}	1.576^{***}	1.336^{***}	1.176^{**}	1.142^{**}	1.289^{***}	1.160^{***}	1.371^{***}
	(0.451)	(0.449)	(0.510)	(0.504)	(0.521)	(0.424)	(0.365)	(0.394)
Leaders^2	-1.232^{***}	-1.218^{***}	-1.020^{**}	-0.845^{*}	-0.817^{*}	-0.966^{**}	-0.887^{***}	-1.042^{***}
	(0.421)	(0.417)	(0.462)	(0.456)	(0.469)	(0.379)	(0.324)	(0.346)
Ln area		-0.083	-0.150	0.100	0.119	0.022	0.027	0.002
		(0.136)	(0.152)	(0.154)	(0.173)	(0.146)	(0.126)	(0.138)
Ln population lag			0.194	-0.375	-0.329	-0.101	-0.152	-0.103
			(0.199)	(0.229)	(0.298)	(0.279)	(0.241)	(0.297)
Ln GDP lag				0.703^{***}	0.666^{***}	-0.037	0.078	-0.023
				(0.156)	(0.220)	(0.316)	(0.261)	(0.329)
Urbanization lag					0.048	0.065	0.026	0.036
					(0.200)	(0.182)	(0.161)	(0.196)
Security expense lag						0.416^{***}	0.276^{**}	0.308^{**}
						(0.142)	(0.123)	(0.132)
Protest time lag							0.234^{***}	0.201^{***}
							(0.071)	(0.072)
Protest spatial lag								0.119
								(0.087)
AIC	798.134	799.764	800.808	783.014	784.956	777.844	768.272	714.694
BIC	817.220	822.667	827.528	813.551	819.310	816.015	810.260	759.455
Log Likelihood	-394.067	-393.882	-393.404	-383.507	-383.478	-378.922	-373.136	-345.347
Num. obs.	336	336	336	336	336	336	336	308
Num. groups: province	28	28	28	28	28	28	28	28

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table A.6: Random Effect Model (28 Provinces and Land Protest)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Intercept	-1.472^{***}	-1.477^{***}	-1.510^{***}	-1.605^{***}	-1.605^{***}	-1.443^{***}	-1.439^{***}	-1.437^{***}
	(0.181)	(0.181)	(0.188)	(0.183)	(0.185)	(0.161)	(0.148)	(0.159)
Leaders	2.231^{***}	2.141^{***}	1.935^{***}	1.749^{**}	1.790^{**}	1.759^{***}	1.668^{***}	1.793^{***}
	(0.631)	(0.634)	(0.700)	(0.692)	(0.723)	(0.578)	(0.578)	(0.594)
Leaders^2	-1.735^{***}	-1.675^{***}	-1.506^{**}	-1.301^{**}	-1.336^{**}	-1.351^{***}	-1.283^{**}	-1.390^{***}
	(0.574)	(0.572)	(0.621)	(0.612)	(0.637)	(0.508)	(0.508)	(0.520)
Ln area		-0.114	-0.184	0.150	0.124	0.007	0.000	-0.019
		(0.167)	(0.197)	(0.195)	(0.232)	(0.196)	(0.193)	(0.202)
Ln population lag			0.175	-0.601^{**}	-0.673	-0.330	-0.398	0.054
			(0.260)	(0.303)	(0.452)	(0.433)	(0.379)	(0.413)
Ln GDP lag				0.975^{***}	1.034^{***}	0.052	0.162	-0.330
				(0.230)	(0.359)	(0.498)	(0.400)	(0.451)
Urbanization lag					-0.067	0.007	-0.047	0.180
					(0.312)	(0.293)	(0.256)	(0.270)
Security expense lag						0.534^{***}	0.462^{***}	0.465^{***}
						(0.196)	(0.171)	(0.174)
Protest time lag							0.116	0.093
							(0.087)	(0.083)
Protest spatial lag								0.229^{*}
								(0.119)
AIC	505.638	507.162	508.704	493.016	494.970	488.844	488.940	467.656
BIC	524.724	530.065	535.424	523.553	529.324	527.015	530.928	512.417
Log Likelihood	-247.819	-247.581	-247.352	-238.508	-238.485	-234.422	-233.470	-221.828
Num. obs.	336	336	336	336	336	336	336	308
Num. groups: province	28	28	28	28	28	28	28	28

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table A.7: Competition Effect Model (Poisson)

	All	Extraction	Land
Intercept	0.192**	−0.614***	−1.444***
	(0.091)	(0.106)	(0.131)
Leaders	1.375***	1.410***	1.756***
	(0.478)	(0.459)	(0.523)
Leaders^2	−1.030***	−1.054***	−1.346***
	(0.392)	(0.378)	(0.431)
Ln area	0.056	−0.008	−0.050
	(0.155)	(0.148)	(0.196)
Ln population lag	−0.539**	0.136	0.224
	(0.225)	(0.251)	(0.313)
Ln GDP lag	0.495**	−0.112	−0.436
	(0.210)	(0.246)	(0.348)
Urbanization lag	−0.087	0.169	0.264
	(0.211)	(0.202)	(0.268)
Security expense lag	0.205***	0.291***	0.468***
	(0.063)	(0.073)	(0.099)
Protest time lag	0.163***	0.104***	0.014
	(0.048)	(0.024)	(0.027)
Protest spatial lag	0.038	0.072	0.191
	(0.073)	(0.135)	(0.170)
AIC	1165.328	772.193	503.133
BIC	1203.647	810.512	541.452
Log Likelihood	-572.664	-376.097	-241.566
Num. obs.	341	341	341

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.
Standard errors clustered at the province level.

Table A.8: Competition Effect Model (Negative Binomial)

	All	Extraction	Land
Intercept	0.159*	−0.619***	−1.451***
	(0.088)	(0.096)	(0.124)
Leaders	1.229***	1.319***	1.685***
	(0.441)	(0.460)	(0.497)
Leaders^2	−0.924**	−0.998***	−1.305***
	(0.363)	(0.377)	(0.412)
Ln area	−0.019	−0.076	−0.110
	(0.135)	(0.117)	(0.179)
Ln population lag	−0.695***	0.036	0.130
	(0.208)	(0.243)	(0.319)
Ln GDP lag	0.664***	−0.068	−0.387
	(0.211)	(0.223)	(0.327)
Urbanization lag	−0.225	0.050	0.155
	(0.172)	(0.179)	(0.256)
Security expense lag	0.194***	0.296***	0.482***
	(0.071)	(0.057)	(0.077)
Protest time lag	0.267***	0.205***	0.098**
	(0.055)	(0.041)	(0.041)
Protest spatial lag	0.082	0.159**	0.256**
	(0.059)	(0.076)	(0.125)
AIC	1081.954	743.154	486.690
BIC	1124.104	785.305	528.841
Log Likelihood	-529.977	-360.577	-232.345
Num. obs.	341	341	341

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.
Standard errors clustered at the province level.

Table A.9: Competition Effect Model (Cross-sectional only)

	All	Extraction	Land
Intercept	2.761***	1.914***	1.051***
	(0.080)	(0.076)	(0.118)
Leaders	1.067***	1.091***	1.155**
	(0.363)	(0.326)	(0.506)
Leaders^2	−0.813**	−0.850***	−0.924**
	(0.331)	(0.293)	(0.457)
Ln area	−0.082	−0.155	−0.195
	(0.125)	(0.119)	(0.183)
Ln population lag	0.198	1.028***	1.221**
	(0.387)	(0.330)	(0.507)
Ln GDP lag	−0.746	−1.507***	−1.891***
	(0.480)	(0.386)	(0.604)
Urbanization lag	0.197	0.535***	0.705**
	(0.240)	(0.198)	(0.305)
Security expense lag	0.890***	0.824***	0.942***
	(0.180)	(0.114)	(0.170)
AIC	224.006	168.545	127.938
BIC	236.912	181.451	140.843
Log Likelihood	-103.003	-75.272	-54.969
Deviance	31.749	37.788	22.340
Num. obs.	31	31	31

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

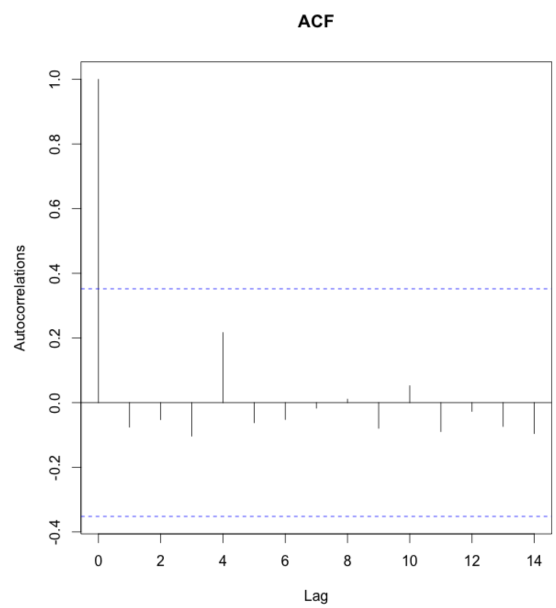
Table A.10: Competition Effect Model (Unscaled Covariates)

	All	All	Extraction	Extraction	Land	Land
Intercept	0.270	0.526	−2.116	−2.207	−2.039	−3.366
	(2.206)	(2.333)	(2.428)	(2.534)	(3.500)	(3.700)
Leaders	0.204***	0.220***	0.201***	0.213***	0.274***	0.279***
	(0.065)	(0.068)	(0.066)	(0.061)	(0.090)	(0.092)
Leaders^2	−0.004**	−0.004***	−0.004**	−0.004***	−0.005***	−0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Ln area	−0.017	−0.051	0.018	0.001	0.005	−0.015
	(0.113)	(0.118)	(0.118)	(0.112)	(0.160)	(0.165)
Ln population lag	−1.039***	−1.022***	−0.116	−0.118	−0.377	0.062
	(0.261)	(0.297)	(0.318)	(0.339)	(0.495)	(0.473)
Ln GDP lag	0.847***	0.817***	−0.032	−0.020	0.046	−0.289
	(0.216)	(0.262)	(0.277)	(0.288)	(0.435)	(0.394)
Urbanization lag	−2.136*	−2.382*	0.431	0.238	0.047	1.190
	(1.195)	(1.354)	(1.203)	(1.289)	(1.933)	(1.781)
Security expense lag	0.003**	0.002	0.005***	0.004**	0.007***	0.006***
	(0.001)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
Protest time lag		0.036		0.129***		0.100
		(0.025)		(0.046)		(0.089)
Protest spatial lag		0.132**		0.112		0.215*
		(0.063)		(0.082)		(0.112)
AIC	1149.002	1070.106	777.844	714.694	488.844	467.656
BIC	1188.191	1116.089	816.015	759.455	527.015	512.417
Log Likelihood	-564.501	-523.053	-378.922	-345.347	-234.422	-221.828
Num. obs.	372	341	336	308	336	308
Num. groups: province	31	31	28	28	28	28

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

I Appendix: Autocorrelation Diagnostics of Protest Data in Each Province

Figure A.5: Autocorrelation Diagnostics



J Appendix: Results Using Active Leaders

Table A.11: Competition Effect Model (Active Leaders)

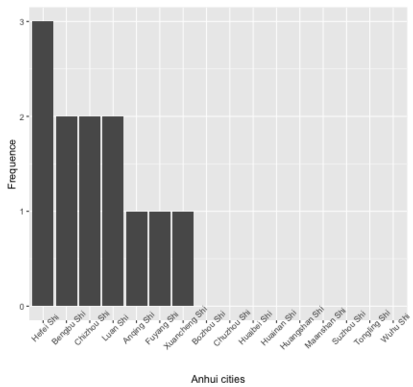
	All	All	Extraction	Extraction	Land	Land
Intercept	0.081	0.052	-0.561^{***}	-0.581^{***}	-1.454^{***}	-1.436^{***}
	(0.108)	(0.111)	(0.114)	(0.114)	(0.163)	(0.168)
Active Leaders	1.367^{***}	1.503^{***}	1.176^{***}	1.336^{***}	1.647^{***}	1.680^{***}
	(0.384)	(0.401)	(0.395)	(0.378)	(0.555)	(0.559)
Active Leaders^2	-1.034^{***}	-1.142^{***}	-0.898^{**}	-1.040^{***}	-1.341^{***}	-1.377^{***}
	(0.364)	(0.374)	(0.365)	(0.341)	(0.514)	(0.513)
Ln area	0.013	-0.022	0.012	-0.005	-0.025	-0.039
	(0.137)	(0.144)	(0.148)	(0.141)	(0.196)	(0.200)
Ln population lag	-0.914^{***}	-0.920^{***}	-0.101	-0.131	-0.336	0.061
	(0.220)	(0.253)	(0.276)	(0.308)	(0.415)	(0.421)
Ln GDP lag	0.996^{***}	0.985^{***}	0.035	0.057	0.137	-0.279
	(0.242)	(0.294)	(0.312)	(0.354)	(0.475)	(0.475)
Urbanization lag	-0.296^{*}	-0.335^{*}	0.069	0.027	-0.004	0.185
	(0.172)	(0.197)	(0.179)	(0.204)	(0.276)	(0.277)
Security expense lag	0.201^{*}	0.124	0.389^{***}	0.296^{**}	0.513^{***}	0.464^{**}
	(0.104)	(0.117)	(0.141)	(0.139)	(0.195)	(0.183)
Protest time lag		0.089		0.195^{**}		0.103
		(0.068)		(0.085)		(0.090)
Protest spatial lag		0.134^{**}		0.114		0.220^{*}
		(0.066)		(0.088)		(0.120)
AIC	1145.426	1065.926	778.022	713.726	489.598	468.290
BIC	1184.615	1111.909	816.193	758.487	527.769	513.051
Log Likelihood	-562.713	-520.963	-379.011	-344.863	-234.799	-222.145
Num. obs.	372	341	336	308	336	308
Num. groups: province	31	31	28	28	28	28

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

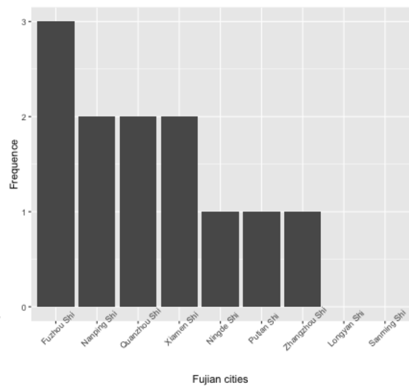
K Appendix: Variation of Protest Within Provinces

Figure A.6: Variation of Protest Within Provinces

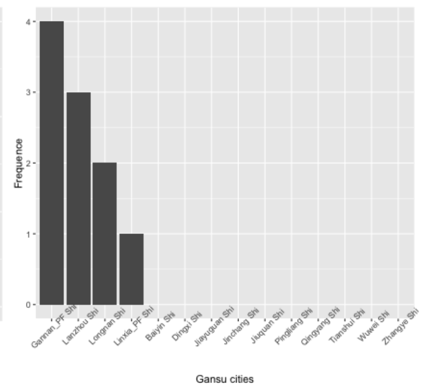
(a) Anhui



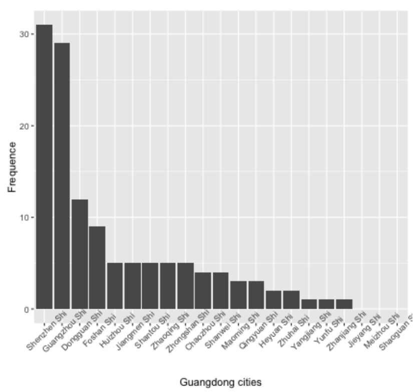
(b) Fujian



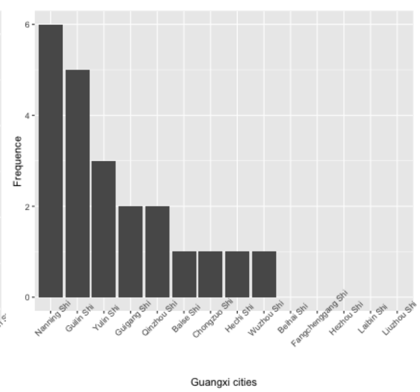
(c) Gansu



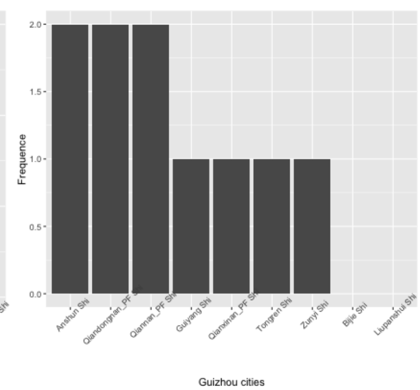
(d) Guangdong



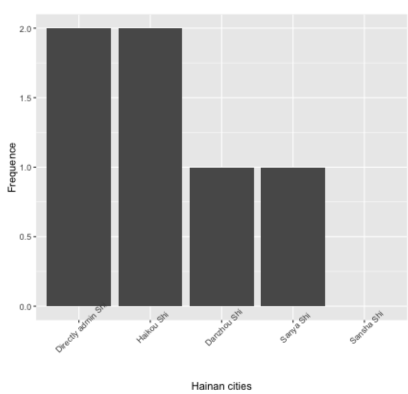
(e) Guangxi



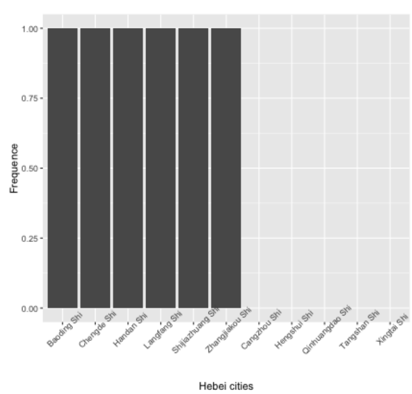
(f) Guizhou



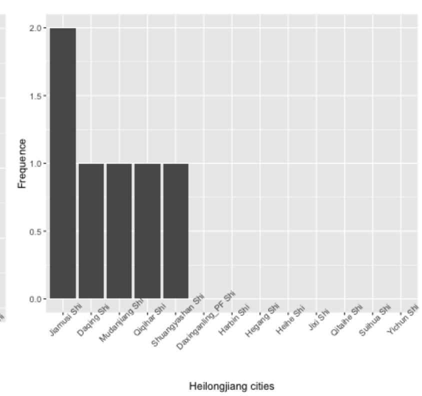
(g) Hainan

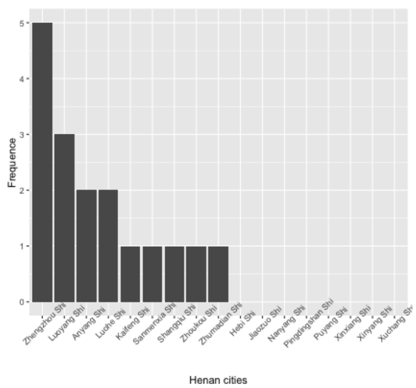


(h) Hebei

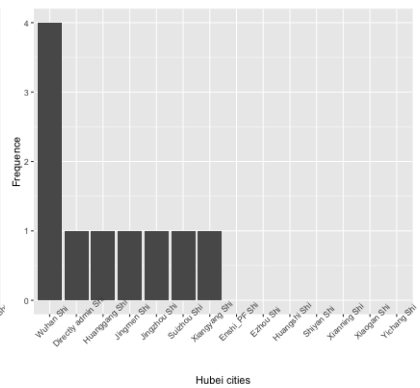


(i) Heilongjiang

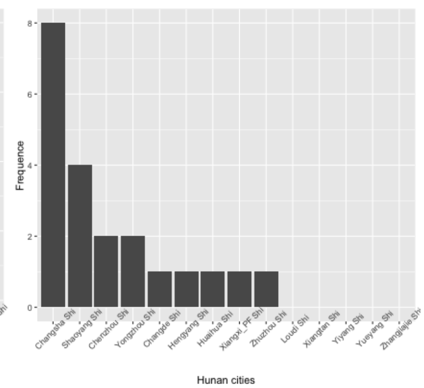




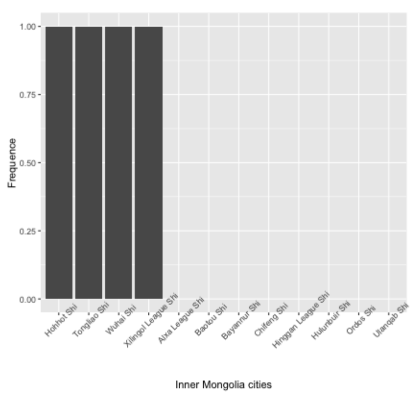
(j) Henan



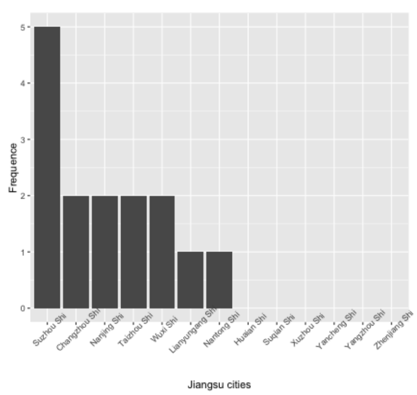
(k) Hubei



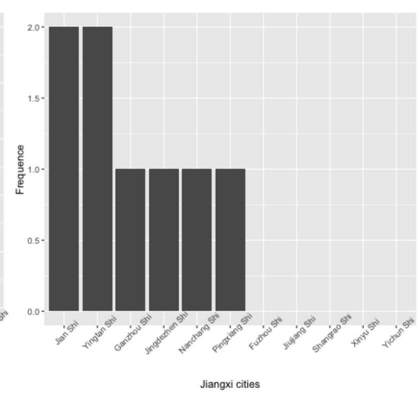
(l) Hunan



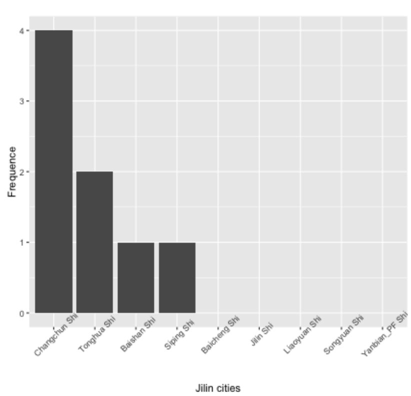
(m) Inner Mongolia



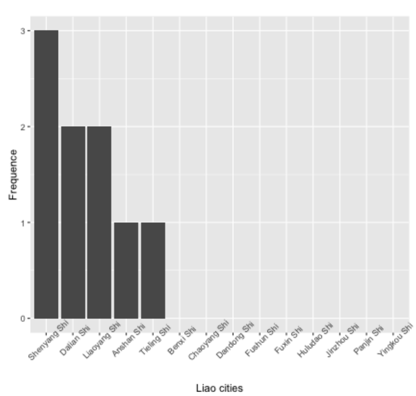
(n) Jiangsu



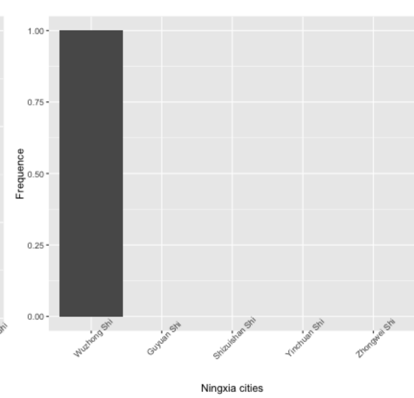
(o) Jiangxi



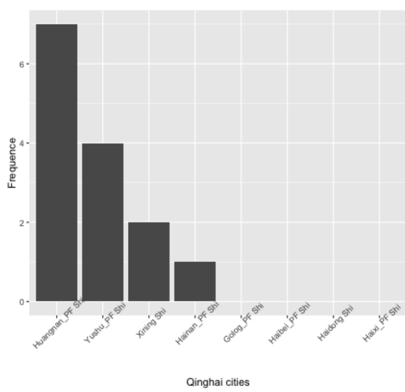
(p) Jilin



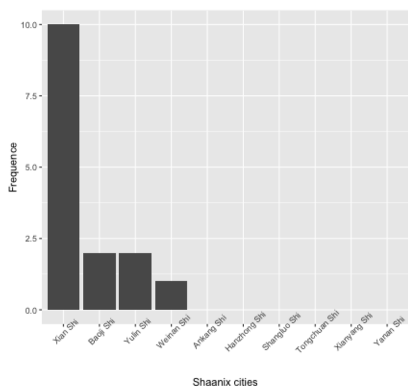
(q) Liaoning



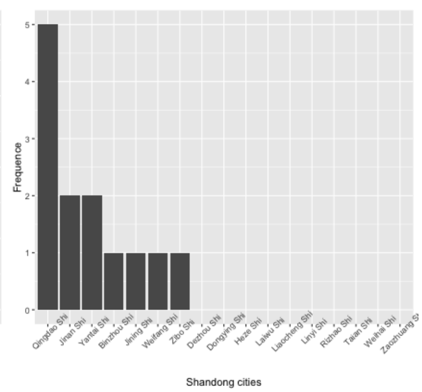
(r) Ningxia



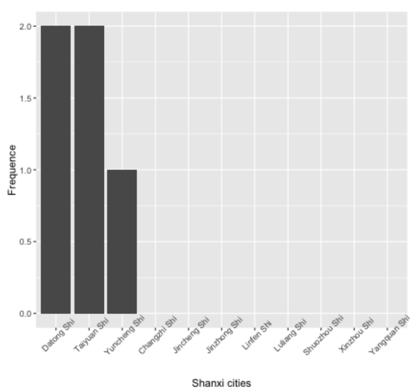
(s) Qinghai



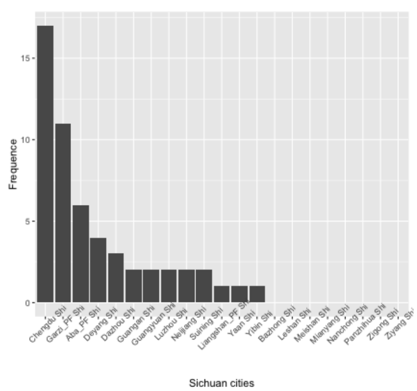
(t) Shaanxi



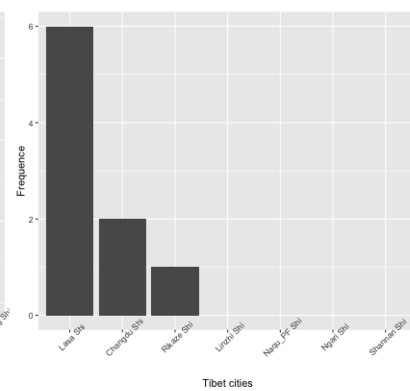
(u) Shandong



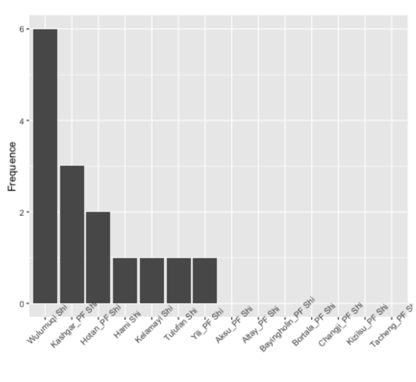
(v) Shanxi



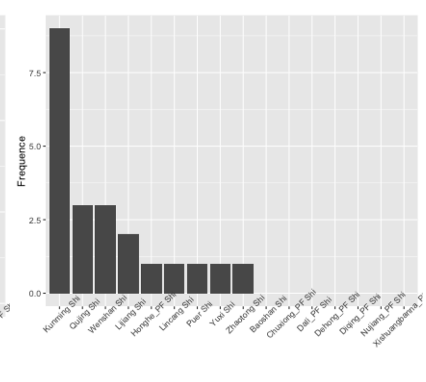
(w) Sichuan



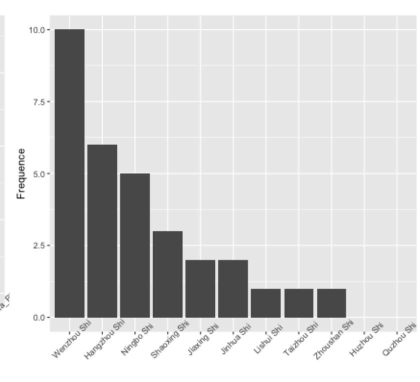
(x) Tibet



(y) Xinjiang



(z) Yunnan



(aa) Zhejiang

L Appendix: Controlling the Temporal Trend of Protests and Potential Censorship in China

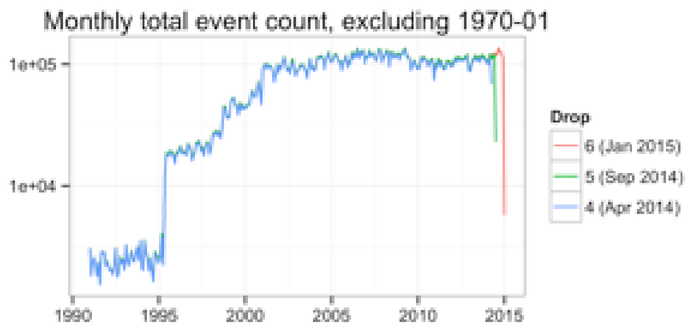
Table A.12: Competition Effect Model (Time Splines)

	All	All	Extraction	Extraction	Land	Land
Intercept	0.437	-0.275	-0.142	-0.785	-2.423^{***}	-3.614^{**}
	(0.447)	(0.794)	(0.527)	(1.004)	(0.938)	(1.635)
Leaders	1.297^{***}	1.362^{***}	1.306^{***}	1.374^{***}	1.750^{***}	1.796^{***}
	(0.428)	(0.446)	(0.409)	(0.390)	(0.563)	(0.591)
Leaders^2	-0.963^{**}	-0.996^{**}	-0.989^{***}	-1.048^{***}	-1.356^{***}	-1.388^{***}
	(0.396)	(0.409)	(0.365)	(0.342)	(0.496)	(0.517)
Ln area	-0.028	-0.054	0.025	-0.008	-0.014	-0.001
	(0.144)	(0.152)	(0.145)	(0.141)	(0.196)	(0.210)
Ln population lag	-0.987^{***}	-1.229^{***}	-0.117	-0.168	-0.382	-0.239
	(0.358)	(0.412)	(0.387)	(0.396)	(0.561)	(0.624)
Ln GDP lag	1.014^{**}	1.286^{**}	-0.091	-0.007	-0.001	-0.100
	(0.439)	(0.499)	(0.453)	(0.431)	(0.603)	(0.683)
Urbanization lag	-0.363^{*}	-0.511^{**}	0.054	-0.007	-0.048	0.056
	(0.202)	(0.242)	(0.203)	(0.222)	(0.310)	(0.355)
Security expense lag	0.281^{***}	0.250^{**}	0.498^{***}	0.373^{***}	0.615^{***}	0.568^{***}
	(0.107)	(0.126)	(0.138)	(0.142)	(0.162)	(0.182)
Time spline	-0.335	0.183	-0.395	0.002	0.391	1.256
	(0.217)	(0.410)	(0.286)	(0.546)	(0.485)	(0.852)
Time spline^2	0.073^{**}	-0.017	0.087^{*}	0.019	-0.030	-0.186
	(0.036)	(0.065)	(0.048)	(0.087)	(0.078)	(0.134)
Time spline^3	-0.004^{**}	0.000	-0.005^{**}	-0.002	0.000	0.008
	(0.002)	(0.003)	(0.002)	(0.004)	(0.004)	(0.006)
Protest time lag		0.054		0.168^{**}		0.061
		(0.069)		(0.074)		(0.086)
Protest spatial lag		0.120		0.068		0.333^{**}
		(0.084)		(0.116)		(0.168)
AIC	1143.020	1070.414	775.916	718.142	488.998	469.654
BIC	1193.966	1127.892	825.538	774.093	538.620	525.605
Log Likelihood	-558.510	-520.207	-374.958	-344.071	-231.499	-219.827
Num. obs.	372	341	336	308	336	308
Num. groups: province	31	31	28	28	28	28

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

M Appendix: ICEWS Event Collections

Figure A.5: ICEWS Event Data Drops



N Hand-coding Protest Event locations

The data generation process for ICEWS generally contains two parts. First, ICEWS uses keywords and phrases (i.e., protest and demonstration) to search and select relevant news articles in their database (i.e., Factiva or Lexis-Nexis). Second, the selected articles went through the automated event coder it uses to extract event features like actors, actions, and locations (Subrahmanian, 2012; Schrodt and Analytics, 2015; Lee, Liu and Ward, 2019). The coded dataset is now publicly available on the Harvard Dataverse[7] but not the original news articles, which are proprietary materials and can only be accessed via Factiva with a license. I had access to the original news articles through my collaboration with Duke University and a license, which allows me to read raw news texts being used to code protest events in my sample (ICEWS China protest events).

By reading the news text myself, I found that ICEWS machine coder is highly accurate in coding locations at the *national* level (Chinese protests are actually geo-referenced to China) but less so at the *subnational* level, which is consistent to the existing finding that cautions the use of machine-coded datasets for micro-level studies due to geo-coding issues (Hammond and Weidmann, 2014).[8] With this caution in mind, I read through the text and hand-coded protest event locations myself. By so doing, I found that ICEWS location coder only achieves 51% accuracy by comparing it to the hand-coded locations in selecting provincial names. This error in coding events subnationally comes largely from (1) mis-referencing protest locations by the reporting locations (e.g. Beijing) that often appear in the first line of the news; (2) assigning city mentions to a wrong provincial administration when city mentions share the same name across different provinces; and (3) failing to capture multiple events and their location mentions contained in the news. These issues are summarized by Lee, Liu and Ward (2019). In their paper, they also develop a supervised machine-learning method to improve the machine-coding accuracy of news articles, but it is beyond the scope of this study. To ensure the location accuracy of ICEWS protest events in China, this study uses human-coded locations as ground truth to avoid machine-coding biases and test the proposed theoretical argument.

The hand-coding process proceeds as follows. In my protest observations, I read through each event and its original news article and then extracted the subnational location names (provinces) for reported protests. Provincial names are rarely missed in news reports since they are the most identifiable administrative units within a country. If there was only one event reported in an article, one location would be coded. If there were multiple protest events reported in an article, all of them would be coded. Table A.13-A.14 are examples showing that mislocated events by machine can be easily corrected by reading through the news content. Since the issue of mislocating events is more concerning at the subnational level than at the national level, correcting mislocated sub-state locations (i.e., provinces) can effectively reduce the threat from the machine-coding error in my analysis.

[7] https://dataverse.harvard.edu/dataverse/icews

[8] Using GDELT, Hammond and Weidmann (2014) indicate that it should be used with caution for geospatial analyses at the subnational level. Its overall correlation with hand-coded data is mediocre and thus research should be careful when approaching to subnational analysis with current machine-coded data.

Table A.13: News Report Examples in ICEWS data

Date: 4/2/2001	StoryID: 26129683
ICEWS location: Beijing city, Beijing	
True location: Zibo city, Shandong province; Suining city, Sichuan province; Xian city, Shaanxi province	
BEIJING, April 2 (AFP) – Several hundred disgruntled workers in eastern China demonstrated in front of local government headquarters on Monday in the latest report of discontent within the country's huge labor force as pressures mount from lay-offs and lagging pay. Workers from the Zibo Chemical Fertilizer Company in Zibo city, Shandong province, marched to the government office to vent their frustration over lack of government support for the company, on which their livelihood depended, a local official told AFP by telephone... Also on Monday, teachers at the Guanyin Elementary and Middle School in southwestern Sichuan province's Suining city went on strike, according to the center. The center said the strike, triggered after the teachers had not received any pay since early this year, left the school's 1,800 students unsupervised... In the third report of worker discontent, the center said 500 workers from a steel factory near Xian, capital of northern Shaanxi province, started blocking an intersection on Friday. The workers, who wanted to draw attention to the factory's failure to pay them for the past 18 months, discontinued the demonstration on Saturday without achieving anything, the center said... [Description of Source: Hong Kong AFP in English—Hong Kong service of the independent French press agency Agence France-Presse]	

Table A.14: News Report Examples in ICEWS

Date: 11/21/2011	StoryID: 20251933
ICEWS location: Lufeng city, Hunan province	
True location: Lufeng city, Guangdong province	
Hong Kong Television Broadcasts Limited [TVB] in Cantonese on 21 November 2011 at 1500 GMT aired video footage by reporter Lo Jo-an from Lufeng city in Guangdong province, where villagers protested against illegal land grabs by "greedy and corrupt officials." The villagers held banners and chanted slogans opposing "corruption" and "dictatorship" as they marched to the city government offices. They dispersed peacefully after being promised an investigation by the government into the alleged illegal land grabs. Over 1,000 Wukan villagers from Lufeng city unfurled banners accusing the local government of illegal land seizures without compensation over the past four to five years. They accused officials of reselling their land to developers for personal profit... [Description of Source: Hong Kong Television Broadcasts Limited in Cantonese (TVB) is non-PRC-owned and airs a large amount of original news and entertainment programming]	

O Out-of-Sample Prediction and Cross Validation

I perform two additional sets of out-of-sample predictions: (1) predictions on each protest category (all extraction, and land protests) and (2) a 5-fold cross-validation test. Table A.15 reports RMSE scores for all provinces in 2013 and 2014, while Table A.16 reports scores for all provinces but without the outlier Guangdong in 2013 and 2014. Prediction results show that the full model (with the competition variable included) generally outperforms the baseline model while slightly underperforming in the extraction protest category in Table A.15 and land protest category in Table A.16. The mixed prediction accuracy rates are likely influenced by a particular pattern of observations in the last two years of the sample (2013 and 2014). A more robust way is to randomly (re)sample the training and test set so prediction results are less likely biased by a particular set of years in the training and the test set, and prediction can be more reliably evaluated. I resort to a commonly used cross-validation approach, which randomly resamples the data into k-folds and evaluates the predictive performance by averaging the RMSE score. I performed a common 5-fold cross-validation, which partitions the data into 5 random subsets and uses 4/5 (80%) of the subsets as the training set and 1/5 (20%) as the test set. The RMSE score is reported in Table A.17. This cross-validation test shows that including the competition variable in the model consistently offers stronger predictions across different protest categories, which provides additional support for my finding.

Table A.15: Out-of-sample Prediction (2013-2014)

	Full model	Base model
All protest	1.98	2.29
Extraction protest	0.95	1.03
Land protest	1.67	1.07

Note: A better prediction in the full model is colored in blue.

Table A.16: Out-of-sample Prediction, no outlier Guangdong (2013-2014)

	Full model	Base model
All protest	1.40	1.41
Extraction protest	0.77	0.74
Land protest	0.65	0.70

Note: A better prediction in the full model is colored in blue.

Table A.17: 5-fold Cross Validation

	Full model	Base model
All protest	2.04	2.23
Extraction protest	1.27	1.36
Land protest	0.83	0.88

Note: A better prediction in the full model is colored in blue.

References

Gallop, Max and Simon Weschle. 2019. "Assessing the impact of non-random measurement error on inference: a sensitivity analysis approach." *Political Science Research and Methods* 7(2):367–384.

Hammond, Jesse and Nils B Weidmann. 2014. "Using machine-coded event data for the micro-level study of political violence." *Research & Politics* 1(2):2053168014539924.

Lee, Sophie J, Howard Liu and Michael D Ward. 2019. "Lost in space: Geolocation in event data." *Political science research and methods* 7(4):871–888.

Luo, Yunhui and Ji Lin. 2003. "The Excessive Competition over Attracting Outside Investment Through Land Leasing: Evidence from Suzhou and Kunshan Developmental Zones." *Reform (Gai Ge, in Chinese)* 6:017.

Schrodt, Philip A and Parus Analytics. 2015. "Event data in forecasting models: Where does it come from, what can it do." *Unpublished manuscript* .

Shen, Xiaodong. 2006. *The Reason and Countermeasure of City Governments' Irregular Competition: Taking Su Xi Chang Region as an Example.* Master's thesis. Retrieved from China National Knowledge Infrastructure(CNKI) network.

Subrahmanian, VS. 2012. *Handbook of computational approaches to counterterrorism.* Springer Science & Business Media.

Tong, Yanqi and Shaohua Lei. 2013. *Social Protest in Contemporary China, 2003-2010: Transitional Pains and Regime Legitimacy.* London, UK: Routledge.

Weidmann, Nils B. 2015. "A Closer Look at Reporting Bias in Conflict Event Data." *American Journal of Political Science* . doi: 10.1111/ajps.12196.