

Appendix Logistic Regression Results

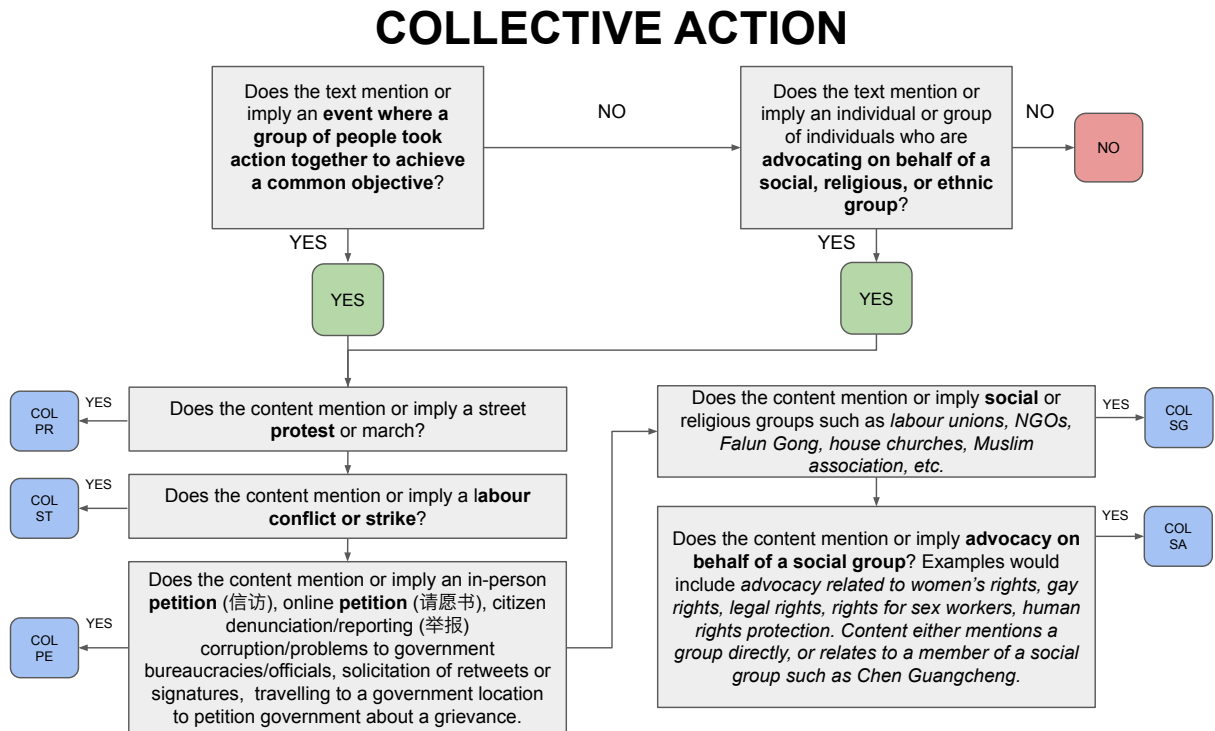
Appendix Table 1: **Logistic Regression Model for “Reporting Up”**

	<i>Dependent variable:</i>
	Report Up
Extreme	1.234*** (0.272)
Govt. Attack	-0.352 (0.243)
User Attributes	1.627*** (0.178)
Rumors	0.083 (0.224)
Retweets	0.630*** (0.206)
Mobilizing	0.967*** (0.228)
Political Humor	-0.744*** (0.253)
Govt. Criticism	-0.007 (0.150)
Local Govt. Corruption	0.225 (0.207)
Collective Action	-0.004 (0.160)
Observations	2,212

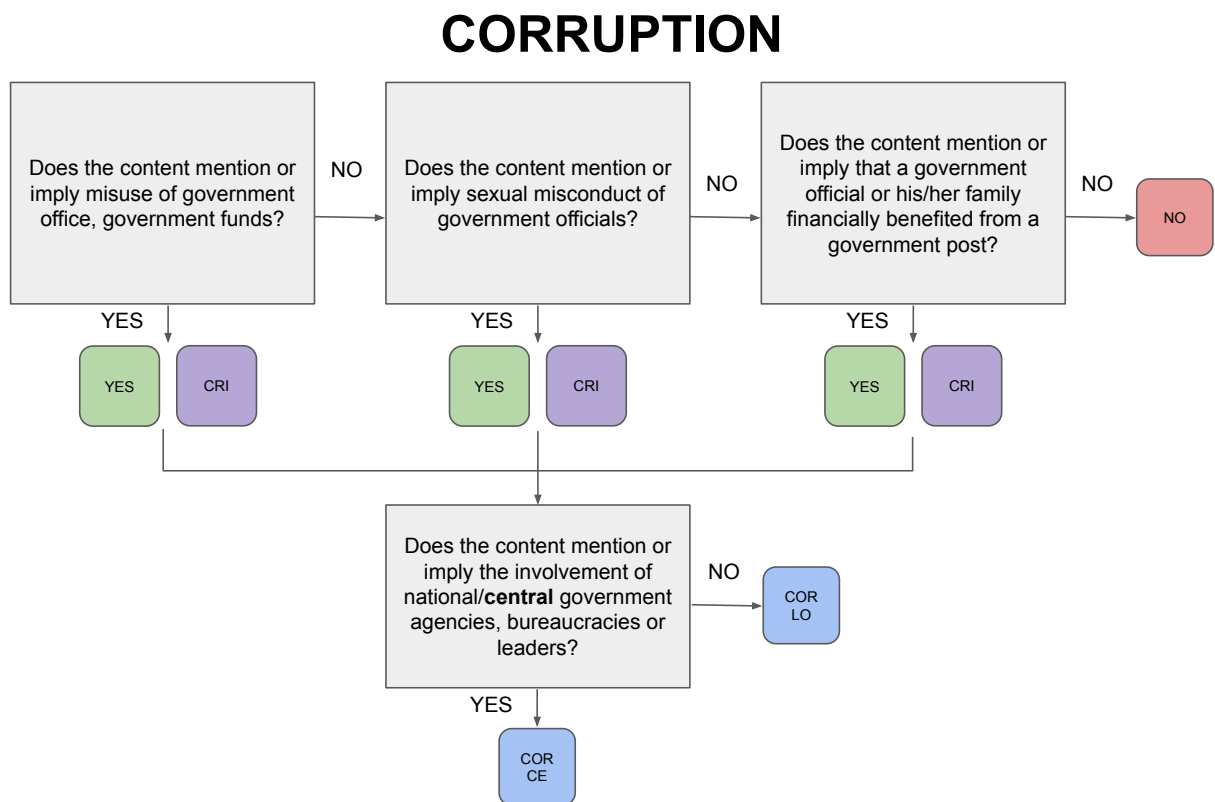
Note: *p<0.1; **p<0.05; ***p<0.01

Content Variables

Appendix Figure 1: Collective Action Coding Diagram



Appendix Figure 2: Corruption Coding Diagram



The instruction text contains a (near) match of the following attributes of users (用户, 账号, 号):

- Small [user] (小[用户,号])
- Big [user] (大[用户,号])
- Big V user – verified user with many fans, V – verified user, VIP (大 V, V, VIP)
- Opinion leaders (意见领袖)
- Famous people (名人)
- Media [users] (媒体)
- Grassroots (草根)
- Important [users] (重要)
- Ordinary [users] (普通)
- Delete for ordinary users and make secret for verified users (普删 V 私)
- [Users] with many fans (粉丝较多的)
- [Users] with around [#] fans (粉丝数[#]左右的)
- Newly registered [accounts/users] (新注册)
- Old accounts (老账户)
- Leftist [users] (左派)
- Rightist [users] (右派)
- [Users] who post memes (段子类)
- “Reincarnated” [users] – users who continually create new accounts after deletion (转世)
- [Users] who continue to post (连续发布的用户)

Extreme

The instruction text contains a (near) match of the following phrases used about content (信息, 内容, 文章号, 言论, 评论):

- Extreme (过激的)
- Excessive/overboard (过分的)
- Vicious/malicious (恶毒的)
- Malicious (恶意的)
- Disgusting/gross (恶性的)
- Disgusting/abominable (恶劣的)
- Inappropriate (不行的)
- Attacking (攻击的)
- Hurling invective (谩骂的)

Attacking the government

The instruction text contains a (near) match of the following phrases used about content the country, the Party, the internet management system, the Party, party-state, country, government, system, social system, leadership, the Central Propaganda Department, Xi Jinping (我国, 我党, 我互联网管理制度, 党, 党政, 国家, 政府, 体制, 社会制度, 领导人, 中宣部, 习近平):

- Attack (攻击)
- Make a negative connection with (负面联系)

- Concerning something negative about (涉及 负面的)

Rumours/fake news

The instruction text contains a (near) match of the following phrases used about content (信息, 内容, 文章号, 言论, 评论):

- Fake (虚假)
- Untrue information (内容不实)
- Spreading fake news (传播虚假消息)
- Spreading rumours (造谣)
- Unconfirmed true or false (不确定真假)
- According [to rumours] (传)
- Refuting rumours (辟谣的)

Retweets

The instruction text contains a (near) match of the following phrases used about content (信息, 内容, 文章号, 言论, 评论):

- Retweets cannot surpass [#] (转发不能过[#])
- Retweets surpass [#] (转发超过[#]的)
- Lots of retweets (转发多的, 转发高的, 转发较高的)
- Every hour surpassing [#] retweets (每小时总量超过[#])
- Pay attention to retweets (注意转发)
- Severely restrict retweets (严禁转载)
- Negative retweets (负面转发)

Inciting

The instruction text contains a (near) match of the following phrases used about content (信息, 内容, 文章号, 言论, 评论):

- Inciting (煽动性)
- Mobilizing (行动性)
- Collective (群体性)
- Call on/appealing to (号召的)
- Cause a fuss, make a scene (闹事)
- Causing a panic (造成恐慌的)

Political humour

The instruction text contains a (near) match of the following phrases used about the country, the Party, the internet management system, the Party, party-state, country, government, system, social system, leadership, the Central Propaganda Department, Xi Jinping (我国, 我党, 我互联网管理制度, 党, 党政, 国家, 政府, 体制, 社会制度, 领导人, 中宣部, 习近平):

- Make fun of (拿 开玩笑)
- Ridicule (调侃)
- Satirize (讽刺的)

- Meme/joke (段子)
- Satirical humour (恶搞)

Intercoder Reliability

These are the results comparing the predictive power of coder 1 to coder 2's decisions using the area under the receiver operator curve (ROC AUC). All categories are above the acceptable measure of .7 and many are above .85, which is highly reliable. We also present another similar measure, Krippendorff's Alpha, which also shows reliability of the relevant categories.

Appendix Table 2: **Intercoder Reliability Measures**

Category	% Agreement	AUC	Krip. Alpha
Coll. action	0.90	0.88	0.76
Corruption	0.95	0.93	0.86
Local corr.	0.95	0.88	0.79
Govt.	0.90	0.84	0.69
Govt. criticism	0.83	0.82	0.65