

# How the Chinese Government Fabricates Social Media Posts for Strategic Distraction, not Engaged Argument<sup>\*†</sup>

Gary King<sup>‡</sup>      Jennifer Pan<sup>§</sup>      Margaret E. Roberts<sup>¶</sup>

June 1, 2016

## Abstract

The Chinese government has long been suspected of hiring as many as 2,000,000 people to surreptitiously insert huge numbers of pseudonymous and other deceptive writings into the stream of real social media posts, as if they were the genuine opinions of ordinary people. Many academics, and most journalists and activists, claim that these so-called “50c Party” posts vociferously argue for the government’s side in political and policy debates. Indeed, the vast majority of posts openly accused on social media of being 50c argue in this way. Yet, almost no systematic empirical evidence exists for this claim, or, more importantly, for the Chinese regime’s strategic objective in pursuing this activity. In the first large scale empirical analysis of this operation, we show how to identify the secretive authors of these posts, the posts written by them, and their content. We estimate that the government fabricates and posts about 448 million social media comments a year. In contrast to prior claims, we show that the Chinese regime’s strategy is to avoid arguing with skeptics of the party and the government, and to not even discuss controversial issues. We infer that the goal of this massive secretive operation is instead to regularly distract the public and change the subject, as most of these posts involve cheerleading for China, the revolutionary history of the Communist Party, or other symbols of the regime. We discuss how these results fit with what is known about the Chinese censorship program, and suggest how they may change our broader theoretical understanding of “common knowledge” and information control in authoritarian regimes.

---

\*Our thanks to Neel Guha, Peter Dyrud, Yingjie Fan, and many others for superb research assistance; Peter Bol, Becky Fair, Chase Harrison, Franziska Keller, Blake Miller, Jean Oi, Samantha Ravich, Andy Walder, Yuhua Wang, Chaodan Zheng, and Yun Zhu for helpful comments; and DARPA (contract W31P4Q-13-C-0055/983-3) and the National Science Foundation (grant 1500086) for research support.

†Authors are listed alphabetically.

‡Albert J. Weatherhead III University Professor, Institute for Quantitative Social Science, 1737 Cambridge Street, Harvard University, Cambridge MA 02138; GaryKing.org, King@Harvard.edu, (617) 500-7570.

§Assistant Professor, Department of Communication, 450 Serra Mall, Building 120, Stanford University, Stanford CA 94304; <http://people.fas.harvard.edu/~jjpan/>, (917) 740-5726.

¶Assistant Professor, Department of Political Science, University of California, San Diego, Social Sciences Building 301, 9500 Gilman Drive, #0521, La Jolla, CA 92093-0521, meroberts@ucsd.edu, MargaretRoberts.net.

# 1 Introduction

Social media in China appears as vibrant and extensive as in any Western country, with more than 1,300 social media companies and websites, and millions of posts authored every day by people all over the country. At the same time, the Chinese regime imposes extensive controls over of the entire system. Which social media companies are prevented from operating in China is easy to see (the so-called “Great Firewall of China”), and the scholarly literature now offers considerable evidence on how and why they censor certain individual social media posts that have appeared on the web or filter them out before appearing. In both cases, the censorship apparatus allows criticism of the regime, its officials, and their policies (which can be useful information for the central government in managing local leaders) but stops discussions that can generate collective action on the ground (King, Pan and Roberts, 2013, 2014).

According to numerous speculations by scholars, activists, journalists, officials in other governments, and participants in social media, the Chinese regime also conducts “astroturfing,” or what we might call “reverse censorship,” surreptitiously posting large numbers of fabricated social media comments, as if they were the genuine opinions of ordinary Chinese people. The people hired for this purpose are known formally as “Internet commentators” (网络评论员) but more widely as “50c party” members (五毛党), so-called because they are rumored to be paid 50 cents (5 jiao, 角, or about US\$0.08) to write and post each comment (Tong and Lei, 2013). (Of course, political parties do not exist in China and so, despite the name, the “50c Party” is not a political party.)

The nearly unanimous view of journalists (and human rights activists) is that 50c party posts strongly argue with and debate against those who criticize the government, its leaders, and their policies. This is also the view of many scholarly publications discussing this activity (and was our view as well, prior to the research reported here). In Section 2, we systematically summarize these views, including a quantitative analysis of social media posts openly accused of being written by 50c party members. Unfortunately, until now, no method has existed for detecting 50c party members, which posts they write, their content, the size of the operation, or why they write them, and so the best anyone could do was

to base claims on intuition, logic, occasional anecdotes, rumors, and leaked government directives.

In this paper, we offer the first systematic empirical evidence for the content of 50c party posts and the government’s strategic objectives.<sup>1</sup> We begin by analyzing an archive of emails leaked from the Internet Propaganda Office of Zhanggong, a district of Ganzhou City in Jiangxi province (Section 3). These emails give explicit details of the work of numerous 50c accounts in this district. Although in the public domain and reported on in the press (e.g., Henochowicz, 2014; Sonnad, 2014), the structure of this archive is complicated, too large to understand by traditional qualitative methods, and in formats (and attachments) far too diverse to make standard methods of automation feasible; as such, it has never before been systematically analyzed, and little of it has been explored. We have developed an approach to analyze this data set and, from it, have extracted more than 43,000 known 50c party posts and their authors.

We first characterize these emails, social media accounts, and known 50c posts via their network and time series structure. Then, in Section 4, we systematically analyze the *content* of the 50c posts in our leaked archive and, in stages, extrapolate to the rest of China. To do this, we first obtain the Sina Weibo accounts of the 50c party members revealed in the leaked emails and obtain from the web many other social media posts they authored. Next, based on measurable features of the authors, their accounts, and the posts in our verified Zhanggong sample (but not their content), we show how to identify 50c party members and their posts throughout China. We then use this methodology to study the content of the posts and finally ascertain the goals behind this massive government program and how it fits with and may reveal broader government strategies.

We validate our 50c party member predictions in Section 5 with a novel sample survey of predicted 50c party members, with several gold standard evaluations (unusual for any survey), to validate the survey and our predictions. In Section 6, we estimate and reveal the size of what turns out to be a massive government operation which writes approximately 448 million posts a year.

---

<sup>1</sup>Because China is a single-party regime, we use the terms “government” and “regime” interchangeably to refer to those in power.

At every stage, our results indicate that the prevailing view of the 50c party is largely incorrect. We show that almost none of the Chinese government’s 50c party posts engage in debate and argument of any kind. They do not step up to defend the government, its leaders, and their policies from criticism, no matter how vitriolic; indeed, they seem to avoid controversial issues entirely. Instead, most of these posts are about cheerleading and positive discussions of valence issues which, we infer, is a strategy designed to actively distract and redirect public attention from ongoing criticism, other grievances, or collective action.

Section 7 gives a unified parsimonious summary of Chinese government internal information control, explores the logic behind it, and shows how these findings may cause scholars to rethink the notion of “common knowledge” in theories of authoritarian politics more generally. Section 8 concludes and then gives a summary of what we might have missed and how scholars can follow up on this work.

## **2 What We Think We Know**

We summarize here views about the 50c party of (1) journalists, (2) academics, and (3) social media participants accused of being 50c party members. The dominant view among (1) and (3) and many in (2) is that 50c party members engage in “hand-to-hand” verbal combat, making specific, directed arguments that support the government, its leaders, and their policies, and opposing arguments to the contrary; they do this by engaging in debate with and criticism of China’s enemies, including those who oppose it inside the country and from abroad. Under (3), we also introduce and validate a scheme to categorize 50c posts on the web; we use it in this section to understand the posts accused of being written by 50c members and then for many other purposes throughout this paper.

Although the difficulties of collecting data on an inherently secret operation means that most prior literature includes “no successful attempts to quantify regime-sponsored commentary in China” (Miller, 2016), the work cited in this section involves considerable effort and creativity on the part of the authors, and even a few clever efforts to guess the values of dependent variables before turning to try to explain or predict the guesses

(Miller, 2016; Han, 2015). Even sophisticated unsupervised statistical techniques have produced “no evidence of large-scale Wumao [50c] activity on Weibo” (Yang, Yang and Wilson, 2015). As most authors acknowledge, claims about the content and extent of 50c party posts rest on little or no empirical basis.

**1. Journalists** The popular press describes 50c members as “undercover pro-government Internet commenters” (Keating, 2011) who “set out to neutralize undesirable public opinion by pushing pro-Party views through chat rooms and Web forums” (Bandurski, 2008). They “shape online public opinion” by labeling “critical opinion leaders as traitors of the country” (Lam, 2012). Prominent dissident Ai Weiwei said “If you oppose the US and Japan [online], you are a member of the 50 cents army” (Strafella and Berg, 2015). 50c party members “combat hostile energy,” defined as posts that “go against socialist core values,” or “are not amenable to the unity of the people.” Such information should be “resolutely resisted, proactively refuted, and eagerly reported to Internet authorities.” (Haley, 2010). Through active engagement of opposition views, they try to “sway public opinion” (Editors, 2016; Ng, 2011), “influence public opinion... pretending to be ordinary citizens and defending or promoting the government’s point of view” (Lam, 2013), or “steer conversations in the right direction” (Editors, 2013). Estimates by journalists of the size of the 50c party is between 500,000 to 2 million (Philipp, 2015).

**2. Academics** Academics have indicated that between 250,000 and 300,000 paid 50c party members write pseudonymous posts as directed by the Chinese government (Freedom House, 2009; Barr, 2012; Greitens, 2013). Because of the absence of systematic scholarly research on the subject, academics express a wider range of possibilities for what 50c party members write about. However, in most cases, their conclusions mirror those of the journalists. They say 50c party members “patrol chatrooms and online forums, posting information favorable to the regime and chastising its critics” (Deibert and Rohozinski, 2010). They “mix control and activism on line... making favorable comments, and generally pushing discussion toward pro-Party lines” (Greitens, 2013). They are an “army of online commentators... promoting the Chinese Communist Party’s line on sen-

sitive subjects” (Bremmer 2010, see also Hassid 2012). They “facilitate state propaganda and defuse crises” (Han, 2015), “post comments favorable towards the government policies” (Tang, Jorba and Jensen, 2012), “defending the government” and “fighting” those who “criticize the government” (Zhang, DiFranzo and Hendler, 2014, p.1889), and, for example, “attack calls for the country to launch a ‘jasmine revolution’” (Bambauer, 2013, p.10).

**3. Social Media Participants** Participants in social media regularly characterize 50c party members by openly accusing others of being members themselves. To systematically characterize their views, we obtained a random sample of 9,911 social media posts from 2010 to 2015 that contain the word “五毛党” (“50c party”). From these data, we drew a sample of 128 posts written by people accused in other posts of being 50c party members.<sup>2</sup>

We then sorted these “accused 50c posts” into one of six categories, using a categorization scheme we will use throughout this paper. With two independent Chinese language coders, and 200 randomly selected posts, we measured the inter-coder reliability of the categorization scheme at 93% agreement (see the Appendix A for details). Two of the categories, comprising 65% of the accused 50c posts, represent the views of academics and journalists, and include (1) Taunting of Foreign countries (which is 29% of this sample) and (2) Argumentative praise or criticism (36% of the sample). Taunting includes favorable comparisons of China compared to other, usually Western, countries and taunting of pro-democracy or pro-West values or opinions. Argumentative praise or criticism involves engaged argument and debate about controversial (non-valence) issues, criticism of opponents of the government, or praise of the leaders.

The categorization scheme also includes (3) Non-argumentative praise or suggestions (22% of the sample) and two categories that everyone agrees are not what 50c party members are writing about, (4) factual Reporting (8%) and (5) Cheerleading (at 5%).

---

<sup>2</sup>Accused 50c party members are distinct from “volunteer 50c members” (自干五; literally “self do 50c” also translated as “bring your own grain”), who express pro-regime or anti-western sentiment online without being paid by the government, the “little red flowers” (小粉红), an unpaid red guard who also attack opponents of the regime online, and the “American 50c” (美分党) who express western democratic values and criticize the Chinese communist regime online. None of these are known to be organized groups.

Non-argumentative praise or suggestions includes discussion of noncontroversial valence issues, such as improving housing or public welfare, or praise of government officials, but does not debate or take opposing viewpoints. (Category (3) does not threaten the regime in any way, and indeed [Chen, Pan and Xu \(2016\)](#) show that local governments openly discuss non-argumentative valence issues with others on government web sites.)

Factual reporting involves descriptions of government programs, events, initiatives, or plans. Cheerleading includes expressions of patriotism, encouragement and motivation, inspirational slogans or quotes, gratefulness, discussions of aspirational figures, cultural references, or celebrations. (Appendix A also includes a sixth “other” or “irrelevant” category, but we remove this so that the percentages from the first five categories add to 100%.)

Thus, social media participants accusing others of being 50c party members agree with scholars and academics that the content of 50c posts is basically antidisestablishmentarianism — arguing with those who oppose with the regime, its leaders, or their policies.

We now go a final step and study the identities of accused (as distinct from actual) 50c party members, which can be difficult because such accusations occur on comment or discussion threads where participants are anonymous. However, by careful and extensive cross-referencing of profile information across multiple platforms, we were able to unearth personal details for a handful of these individuals. Their backgrounds vary greatly, but in each case, it seems obvious that they are highly unlikely to be real 50c party members. For example, among those accused of being 50c party members include Zhou Xiaoping (周小平), a blogger well known for his anti-west and nationalist sentiment, and He Jiawei (何家), a blogger known for critiques of the Chinese government who posts on Boxun, a site hosted outside of China devoted to covering topics such as Chinese government human rights abuses. Other well known figures accused of being 50c include Lin Yifu (林毅夫), a Peking University professor who was Chief Economist and Senior Vice President of the World Bank from 2008 to 2012. In none of these cases are these people likely to be actually 50c party members. However, those accused of be-

ing 50c party members also include figures not connected to politics such as, in our data, a comedian, a lawyer, and a marketing executive.

It appears that the evidence base of those accusing others of being 50c party members is no better than that of academics or journalists. Although the prior beliefs of all three groups about the content of 50c party posts are almost the same, little evidence supports their claims.

### **3 Leaked Internet Propaganda Office Communications**

The problem in the literature has been that “detecting the Wumao [50c party] is difficult because there is no ground truth information about them” (Yang, Yang and Wilson, 2015). We are fortunate to be able to change this situation. In December 2014, anonymous blogger “Xiaolan” (<https://xiaolan.me/>) released an archive of all 2013 and 2014 emails to, and some from, the account of the Internet Propaganda Office (网宣办), a branch of the propaganda department, of Zhanggong district. Zhanggong district is a country-level administrative unit (with a population in 2013 of 468,461) that is part of the moderate sized Ganzhou City, in Jiangxi province. The emails reported activities of Internet commentators, including numerous 50c posts from workers claiming credit for completing their assignments, and many other communications. The hack was widely reported and the archive of emails have been publicly available since (e.g., Henochowicz, 2014).

The archive’s large size, complicated structure, numerous attachments, diverse document formats (screen shots, Word, Excel, Powerpoint, raw text, text as part of other emails, etc.), multiple email storage formats, and many links to outside information has made digesting much of it impossible either for individuals reading and coding by hand or for existing methods of automated text analysis. Journalists managed to pull out a few examples to write newspaper articles, but no systematic analysis has been conducted of these data.

To systematize this richly informative (and essentially qualitative) data source, we developed and applied a variety of methods and procedures, from large scale hand coding,



to specially tuned and adapted methods of named entity recognition, to methods of automated text analysis and extraction. Because of the considerable effort and resources necessary, we will make structured and easy-to-access forms of these data publicly available so that others may follow up. Full replication information for all analyses in this paper will be available as well.

From this work, we identified 2,341 emails sent from 2/11/2013 to 11/28/2014. Of these, 1,245 contained the text of one or (usually many) more 50c posts. In all, we harvested from these emails and their attachments 43,797 known 50c posts that form a basis for our subsequent analyses and, as a training set, help identify other 50c posts. (Although we have the name, direct contact information, and often photographs of many of the people discussed in this paper, we have no academic reason to make this information more public than it already is; we therefore do not do so.) We conduct rigorous evaluations of our claims to the next section. For now, we characterize the content with three separate descriptive analyses.

First, we portray the overall structure of communications in these emails with the network diagram in Figure 1. Each circle is a specific email account and each line denotes where one or more emails was sent from and to. The large flower-like shape at the bottom are 50c party members sending in copies of their posts to the Zhanggong district Internet Propaganda Office (章贡区网宣办), claiming credit for completing their assignments. This office then reports up to other offices (see the lines out from the center of the flower shape), These include the speaker of Zhanggong People's Court News office (江西省赣州市章贡区人民法院新闻发言人), and District Party Office Information Department (区委办信息科).

Second, although the scholarly literature describes 50c party members as ordinary citizens hired for very low wages, we found instead that almost all in our sample are government employees.<sup>3</sup> Of the 43,797 posts, only 291 were made by individuals or groups we could not identify (the content of these posts were very similar to those we could identify). The remaining 99.3% were contributed by one of over 200 government agencies

---

<sup>3</sup>Han (2015), who seems to have gotten it right, argued that internet commentators are often public servants though some may be specially recruited.

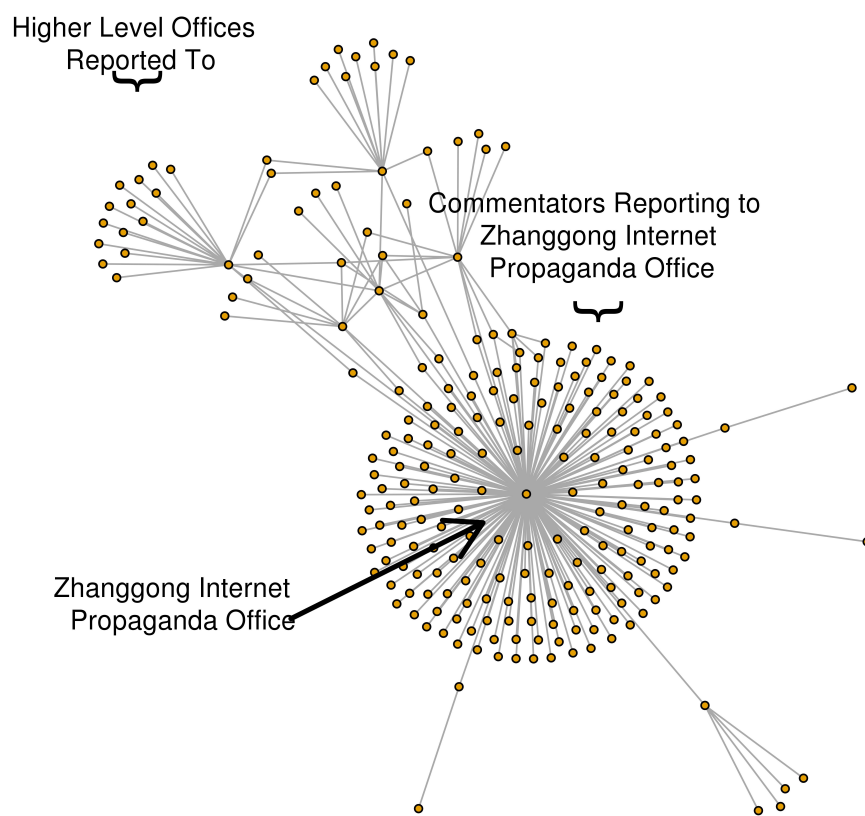


Figure 1: Network Structure of Leaked Email Correspondents. Circles are email correspondents, edges (lines) indicate email correspondence. Most of the correspondence is in toward the center of the flower-like structure (to the Zhanggong Internet Propaganda Office and then out from that office to higher level offices).

throughout the Chinese regime’s matrix organizational structure (of geographic representation by functional area) in Zhanggong district. This included 9,159 posts (20.1% of the 43,797 total) made directly by the Zhanggong Internet Propaganda Office, 2,343 (5.3%) by the Zhanggong district Bureau of Commerce (区 商务局), 1,672 by Shuixi Township (水西镇, one of several townships in Zhanggong), 1,620 by Nanwai sub-district (南外街道 (one of several sub-districts in Zhanggong)). Others come from functional bureaus in Zhanggong district (e.g., 体育局 Sports Bureau, 人保局 Bureau of Human Resources and Social Security, 地税局 Bureau of Taxation, 法院 Zhanggong district court), the government offices of Zhanggong’s sub-districts and townships (e.g., 沙河镇 Shahe Township,

赣江街道 Ganjiang sub-district), functional departments in each sub-district or township (水西镇党政办 Shuixi Township Party Office), and administrative offices of neighborhoods and villages in Zhanggong's townships and sub-districts (e.g., 南外街道东阳山社区 Dongyang Shan neighborhood of the Nanwai sub-district, 水西镇何乐村 Hele village of the Shuixi sub-district).

Of the 50c posts in this archive, 29.98% did not contain a URL or a description of the site where the content was posted. Of the remainder, 53.54% of the 50c posts were comments on government sites (GanzhouWeb, Newskj, DajiangWeb, JidanWeb, JiangxiWeb, CCTVWeb, RenminWeb, JiujiangWeb, QiangGouWeb), and 46.67% were on commercial sites. Of the 50c posts on commercial sites, 54.06% were on Sina Weibo, 32.05% on Tencent Weibo, 10.73% on Baidu Tieba, and 2.69% on Tencent QZone, with the rest in the long tail receiving less than 1% each.

We also found no evidence that 50c party members were actually paid fifty cents or any other piecemeal amount. Indeed, no evidence exists that the authors of 50c posts are even paid extra for this work. We cannot be sure of current practices in the absence of evidence but, given that they already hold government and Chinese Communist Party (CCP) jobs, we would guess this activity is a requirement of their existing job or at least rewarded in performance reviews.

Finally, we offer a first look at the content of the 43,797 posts from the 50c party that we unearthed. We do this by plotting in Figure 2 a daily time series of counts of these posts. The most important finding in this graph is that the posts are far from uniformly distributed, but instead are highly focused into distinct volume bursts. This suggests a high level of coordination on the part of the government. Indeed, often the most influential patterns on social media are the bursts that occur naturally when discussions go viral. The government's manufactured bursts mirror these influential naturally occurring patterns. Bursts are also much more likely to be effective at accomplishing specific goals than a strategy of randomly scattering government posts in the ocean of real social media.

Although we will conduct rigorous, quantitative analyses of the content of 50c posts in the sections to follow, we also provide a feel for the general content and time line of

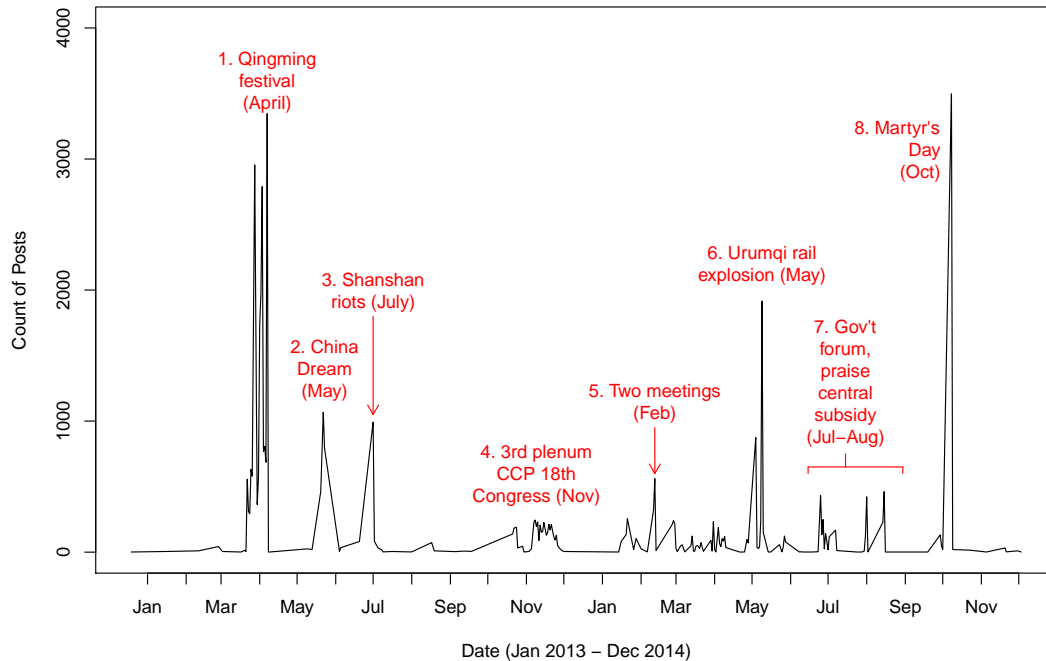


Figure 2: Time series of 43,797 known 50c social media posts, with qualitative summaries of the content of volume bursts.

these posts by labeling the largest volume bursts in this set. The labels are brief qualitative summaries we chose from reading numerous posts.

As with most social media time series plots like this in most countries, regardless of classification, the 50c posts in this collection cluster in volume bursts. We thus describe each volume burst (with numerical labels corresponding to those in the figure). As can be seen throughout, this is the first indication that the overwhelming focus of these posts is on cheerleading and distraction, rather than engaged argumentation and debate. The timing some of these bursts may even suggest a possible government strategy of following collective action events with posts to distract the public with an interesting, but innocuous and unrelated topic.

1. Qingming (Tomb Sweeping Day): over 18,000 posts about veterans, martyrs, how glorious or heroic they are, and how they sacrificed for China.
2. China Dream: Over 1,800 posts about President Xi Jinping’s “China Dream”. Potentially a reaction to the April 2013 People’s Daily piece instructing municipal

governments to carry out “China Dream” propaganda campaigns (see <http://j.mp/chinadream>).

3. Shanshan Riots: 1,100 posts, immediately following Shanshan riots in Xinjiang. At 5:30pm Zhanggong county sent an email to itself (probably BCCing many others), highlighting three popular posts about Xinjiang, and saying this was a terrorist incident. At 8:00pm on the same day, Zhanggong sent an email to Ganzhou City to which it reports with hundreds of 50c posts about China Dream, local economic development, etc.
4. 18th Party Congress, 3rd Plenum: Over 3,400 posts related to the 3rd plenary session of the Chinese Communist Party’s 18th Congress, which discussed plans for deepening structural reform.
5. “Two Meetings”: Over 1,200 posts about Ganzhou’s People’s Congress and Political Consultative Committee meetings, and policies to be discussed at the two meetings, including factual reporting of environmental issues, one child policy, rural issues, growth and development.
6. Early May Burst: 3,500 posts, about a variety of topics, e.g., mass line, two meetings, people’s livelihood, good governance. Immediately followed the Urumqi Railway Explosion, but no connection other than timing is apparent.
7. Over 2,600 posts celebrating the second anniversary of “Central Soviet Areas Development policy” (若干意见), subsidies from the central government to promote the development of region where the original CCP bases were located (including the region where Zhanggong is located); at the same time, the local government held an online Q&A session for citizens.
8. Martyr’s Day: 3,500 posts about martyrs and the new Martyr’s Day holiday, celebrating heroes of the state.

We now turn to a more systematic analysis of these posts, their accounts, and others like them beyond Zhanggong.

## 4 Content of 50c Posts

We now reveal the content of our 50c party posts from across China by estimating the distribution of these posts across the five main content categories introduced in Section 2 (with details in Appendix A). We do this in five separate analyses and data sets which successively expand the target set of 50c party posts to larger and larger areas across the country, as much as possible making ourselves vulnerable to being proven wrong. Through each of these analyses, we find very similar results, with 50c party posts largely comprised of cheerleading and distraction rather than engaged argument.

**1. Leaked 50c Posts** We first analyze the 43,797 50c social media posts we harvested from numerous authors and different social media sites in the leaked archive from Zhang-gong. To study these data, we first hand code a random sample of 200 posts into our six categories, with high inter-coder reliability.

One result is immediately apparent: The number of posts from this sample that fall in the categories “taunting of foreign countries” or “argumentative praise or criticism” is exactly *zero*. This is an important surprise, as it is essentially the opposite of the nearly unanimous views espoused by scholars, journalists, activists, and social media participants. This result would be highly unlikely to have resulted from (Binomial) sampling error if the true share of the full set were even as large as a few percentage points (at 5%, which would still be a major surprise, the probability of seeing the sample we obtained is essentially zero). To push even farther, we did extensive searches and reading among the remaining posts and finally found a few that fit this category (see the examples in Categories 2 and 3 in Appendix A), but the overall result is that 50c party posts are extremely rare in these categories.

We thus infer that the leaked posts contain very little Taunting of Foreign Countries or Argumentative Praise or Criticism, but we go ahead and verify by formally estimating all the category proportions in the entire set of posts. Using a text analytic method known colloquially as “ReadMe” (named for the open source software that implements it), we estimate the category proportions directly, without having to classify each post into

a category (Hopkins and King, 2010). This is fortunate, since individual classifiers that manage to achieve high (but imperfect) levels of the percent correctly classified, may still generate biased estimates of the category proportions. For example, an estimate indicating that zero country dyad-years, since WWII, were at war achieves a predictive accuracy of about 99.9%, but aggregating these classifications yields an obviously biased (and useless) estimate of the prevalence of war. In contrast, ReadMe does not give individual classifications, but it has been proven to give approximately unbiased and consistent estimates of the category proportions, which here is the relevant quantity of interest. The other advantage of ReadMe in this context is that its statistical assumptions are met exactly by our sampling procedures.

The estimated proportions of 50c posts by category for all data sets appear in Figure 3; the results for our first data set (of all posts found in the leaked emails in Zhanggong) are represented by a histogram, formed by the set of solid disks (●) for the point estimate, and solid line for the confidence interval, for each of the categories. Other results, to be described below in order from left to right within each category, also appear in the same graph.

The categories in Figure 3 are arranged so that the two emphasized in the literature appear on the left, and our main empirical results on the right. For this analysis, the results indicate that approximately 80% fall within the Cheerleading category, 13% in Non-argumentative praise or suggestions, and only tiny amounts in the other categories, including nearly zero in Argumentative praise or criticism and Taunting of foreign countries. Clearly, these results clearly indicate that 50c posts are about cheerleading, not argumentation.

**2. Posts from Leaked 50c Weibo Accounts** One possibility we now consider is whether 50c party members differentially reported cheerleading posts back to the Propaganda Department, even though they posted about topics at the behest of the regime from other categories as well. To study this question, we construct a second data set by first identifying all Weibo social media accounts revealed in the leaked email archive. We choose Weibo because it is the most widely used social media site that enables mass distribution,

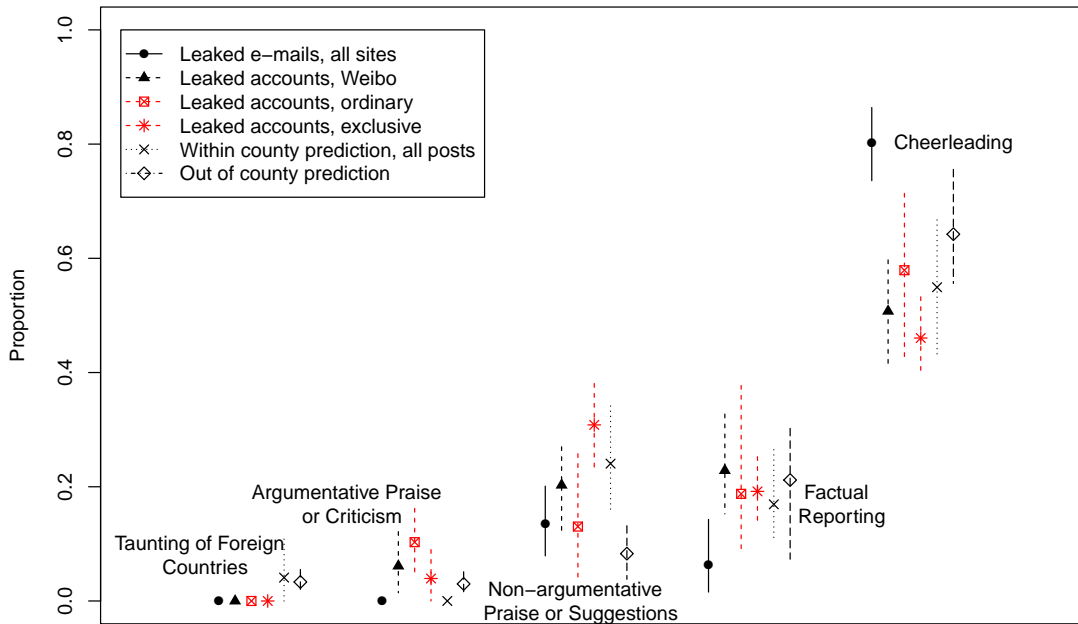


Figure 3: Content of Leaked and Inferred 50c Posts, by substantive category (with details in Appendix A) and analysis (given in the legend)

and we were able to obtain access in the manner we needed it. We then found these accounts on the web and kept all 498 Weibo accounts that made at least one post. Finally, we downloaded all social media posts from these accounts, yielding a set of 167,971 known — but not previously leaked — posts from 50c accounts.

We drew a random sample (stratified by account) of 500 of these 167,971 social media posts, and coded them into our categories as a training set. In this randomly selected training set, like the last, we find no evidence of taunting of foreign countries, although we did find a handful of posts in the category of Argumentative Praise or Criticism, constituting only 3% of the posts. As above, we then used (a stratified sample and) ReadMe to estimate the five category proportions for the set of all posts. The results, reported in the second bar of the histogram in Figure 3, are very similar to that from the first data set. The point estimates (portrayed as solid triangles, ▲, with confidence intervals as dashed lines) indicate that again the bulk of 50c posts from leaked accounts are Cheerleading (51%), 22% in Factual Reporting, 20% in the category Non-argumentative praise or suggestions,



and only 0.4% in the category Argumentative praise or criticism.

**3. Partitioning Leaked Accounts for Extrapolation** We designed our third analysis to further explore the leaked data as well as to prepare the ground for extrapolation. The key idea here is to partition accounts into those easy to identify outside the leaked archive (which we do for Zhanggong in Analysis 4 and in other counties in Analysis 5) and those more difficult to identify. We develop an algorithm to distinguish these two account types and then show that we only need to extrapolate the first type because they post the same types of content.

We first study the structure of the 498 known 50c Weibo accounts and their 167,971 social media posts, and found two fundamentally different types of accounts. In both, we often find many commercial posts, which fall in our “other” category (see the Appendix). Since we remove and condition on this “other” category for all analyses, we do not define account types on this basis either. Thus, the first type of account, which we call *ordinary*, are used by apparently ordinary people in China to post about their children, funny videos, commercial advertisements, sports teams, pop stars, and many other subjects. Embedded within the stream of these posts are others which these authors indicate in their communication with the Propaganda Department to be 50c party posts. The second, which we call *exclusive* accounts, are (aside from commercial posts) almost exclusively devoted to 50c posts. Near as we can tell, via extensive cross-checking with external data sources, ordinary accounts are genuine, registered in the name of the government employee posting on it, whereas exclusive accounts are pseudonymous, designed solely to fool those who see it. In both cases, the 50c posts on these accounts are those directed by the government rather than necessarily reflecting the opinions of ordinary people.

Distinguishing between ordinary and exclusive accounts in our leaked archive is easy (the number of real 50c posts reported to the Propaganda Department, as a proportion of all posts on the account, is a direct measure), but our goal is to extrapolate to other counties where we have no known 50c posts. Thus, we need a formal partitioning algorithm to sort accounts into these two categories without needing the inside information we have from our extraordinary leaked data. Moreover, since our goal is to determine the content of 50c

posts, we must be able to discern whether an account was written by a 50c party member without using the text of the posts.

To develop this partitioning algorithm, we followed the logic of “Bayesian falling rule list” methodology, which are accurate and also highly interpretable (Letham et al., 2015). The interpretability also enabled us to combine qualitative knowledge with modern machine learning, as well as to make choices that were much easier to apply outside of Zhanggong. With this approach as a guide, we found that only two rules were sufficient to partition our 498 50c accounts into exclusive and ordinary: First, we obtain candidate 50c accounts by collecting all accounts that comment or forward on any post on the Zhanggong government weibo account (<http://weibo.com/u/3880516376>). Second, we narrow this to accounts with 10 or fewer followers. The result is a set of exclusive accounts. These two simple, interpretable rules are highly plausible and consistent with what is known about social media. After all, accounts that engage with government websites and have no more than a handful of followers were surely created or at least used for a very narrow purpose.

We now show that the 50c posts appearing on exclusive and ordinary accounts have essentially the same types of content, where we can verify both. To do this, we apply our partitioning algorithm to the set of 498 known 50c accounts from our archive and then compare the content of ordinary and exclusive accounts. We find that 202 (41%) are exclusive accounts and the remaining 296 (59%) are ordinary accounts. This partition of the data is neither right nor wrong (and so statistics like “percent correctly classified” do not apply), but it is useful only to the extent that using only the exclusive posts causes no bias. Thus, we estimate and compare the distribution of posts within the ordinary and exclusive account types across our five content categories. To do this, we apply ReadMe within each partition and compare the results.

Fortunately, the results are very close to each other, and (as a result) to the overall results we presented above. This implies that bias is unlikely to be induced by narrowing our search outside our leaked archive to exclusive accounts. Point estimates for the category proportions appear in Figure 3 (marked as  $\square$  and  $\times$ , in red). For both, the bulk of 50c

posts appear in the Cheerleading category (46% for exclusive accounts and 58% for ordinary accounts). In contrast, the sum of taunting of foreign countries and of argumentative praise or criticism is very small (4% for exclusive and 10% for ordinary).

**4. Unleaked 50c Posts in Zhanggong** We now use the results of Analysis 3 and expand our study of 50c posts beyond the leaked archive. In this section, we focus on previously unidentified 50c posts in Zhanggong. To do this, we choose exclusive accounts (by applying the two rules from the previous section). With this procedure, we find 1,031 accounts, of which 829 accounts are not mentioned in our leaked archive. We then find and scrape all 22,702 social media posts available from the front page of each of these accounts. Each front page has up to 45 separate posts. We then analyze these posts with ReadMe, as above.

Results from this analysis appear in Figure 3 (with point estimates represented by  $\times$ ). The result again is very similar to previous analyses: 55% of the posts made on these accounts engaged in Cheerleading, 17% engaged in Factual Reporting, 24% engaged in Non-argumentative praise and suggestions, and about 4% in Taunting of foreign countries and essentially zero in Argumentative praise or criticism.

**5. Unleaked 50c Posts in Counties with County Government Weibo Accounts** We now extrapolate to counties across China. We do this by starting with all 2,862 counties (and county-level divisions). We then take as our target of inference 50c behavior in 1,338 of these counties which are structured same way as Zhanggong, with a Propaganda Department that has a public web site. We then drew a simple random sample of 100 of these counties and identified all exclusive accounts and a sample of their social media posts.<sup>4</sup>

To be more specific, for each county government Weibo account, we collected all 151,110 posts, randomly sampled up to 200 posts of these, identify all outside Weibo accounts that commented on or forwarded any one, downloaded all meta-data from those

---

<sup>4</sup>Many of the remaining 1,524 counties have Weibo accounts run by government bureaus and agencies (such as public security department, civil affairs department), but not by the county-government. Our informal study of these counties revealed no systematic differences from those we studied, but following up with systematic study in these counties would be a good topic for future research.

accounts, and subsetting to those with 10 or fewer followers. We then download the first page, comprising up to 45 social media posts from each account, as our candidate 50c posts.

Figure 3 provides our results (with point estimates represented as a diamond,  $\diamond$ ). Again, we find very similar results, highly focused on cheerleading and distraction rather than argumentation and criticism: 64% of the posts made on these accounts are categorized as Cheerleading, 21% in Factual Reporting, 8% Non-argumentative praise and suggestions, 0.2% in Taunting of foreign countries, and only 1% in Argumentative praise or criticism.

## 5 Verification by Direct Survey

We now attempt to go an extra step to verify the accuracy of our extrapolation in Section 4 to predicted 50c party members across China. To do this, we take the unusual step, in this context, of conducting a sample survey of predicted 50c party members, along with gold standard elements designed even to validate this method of validation.<sup>5</sup>

**1. Design** We began by creating a large number of pseudonymous social media accounts. This required many research assistants and volunteers, having a presence on the ground in China at many locations across the country, among many other logistically challenging complications. We conducted the survey via “direct messaging” on Sina Weibo, which enables private communication from one account to another. With IRB permission, we do not identify ourselves as researchers and instead pose, like our respondents, as ordinary citizens. Since information in our archive appears to indicate that government monitoring of 50c party member activities occurs only through voluntary self-reporting up the chain of command, our survey questions and the responses are effectively anonymous, which are conditions that have been shown to make respondents more sincere in

---

<sup>5</sup>We had full IRB approval for our study (although we cannot make public our exact question wording in Chinese, sample size, certain statistics, or the original data). We also added our own additional ethics rules not required by the IRB designed to further protect the identities of our respondents and to keep our large research team safe. Our rules followed the principle articulated in King, Pan and Roberts (2014, fn.20) of trying to avoid influencing the system we were studying, which has the added advantage of reducing the chance for bias.

responding to sensitive questions (Tourangeau, Conrad and Couper, 2013).

We drew a random sample of social media accounts we predicted in Section 4 to be 50c and asked each whether he or she was indeed a 50c party member (in a special manner described below). Of course, the difficulties of interpreting these answers is complicated by the fact that our survey respondents are conducting surreptitious operations on behalf of the Chinese government designed to fool participants in social media into thinking that they are ordinary citizens, and we are asking them about this very activity. In most cases, the government is also their employer, and so they have ample incentives to not comply with our requests, or to not comply sincerely.

We address these uncertainties with two entire additional surveys designed to provide internal checks on our results, as well as a carefully worded survey question in our anonymous survey context. In most surveys, researchers are left trusting the answer, perhaps after a stage of pretesting or cognitive debriefing. In our survey, we are in the unusual position of being able to go further by offering a gold standard validation, where, for some respondents, we know the outcome to the question we are posing. That is, we ask the same question of a random sample of *known* 50c party members from our Zhanggong leaked archive. If the results of our survey of predicted 50c party members give similar results as this survey, then we should have more confidence in the results.

We also offer a third entire survey that approximates the opposite gold standard by asking those known *not* to be 50c party members. To do this, we draw a random sample from Weibo accounts across China among those who do not engage with government Weibo accounts and have more than 10 followers. Our results would be confirmed if the percent who say they are 50c in this sample are significantly lower than those who acknowledge being 50c in our predicted 50c sample. A tiny fraction of these accounts may actually be 50c, but that would merely bias the results against the test of our hypothesis of the difference in means from our set of predicted 50c members.

The final way we reduce uncertainty is in the design of our survey question. We followed best practices in designing survey questions about sensitive topics, including adjusting the perceived social environment (Näher and Krumpal, 2012) and using familiar

language and positive “loading” of sensitive questions (Groves et al., 2011). We also studied a large volume of social media interactions, both via automated means (King, Lam and Roberts, 2016) and by direct reading, and found a way within the cultural context to ask the question so that it would be more likely to elicit a sincere answer. We also pre-tested our survey on an independent sample. Although preserving the confidentiality of our respondents and research team makes us unable to share the exact text of our question, we report here a similar version in English, which will also enable us to explain its features:

I saw your comment, it’s really inspiring, I want to ask, do you have any public opinion guidance management, or online commenting experience?

To avoid interfering or influencing the system we are studying, and to avoid putting our respondent in an uncomfortable position, the question discusses online propaganda in positive terms. We use the term “opinion guidance management” and “online commenting,” which we learned from our leaked archive is the way the government discusses these tasks. We avoid terms like 50c or “Water Army” which has negative connotations for some. Instead of asking someone to “out” themselves as a 50c party member, we ask for advice on where the person learned to write in such a motivating, inspiring manner, thus avoiding generating defensiveness on part of the respondent.

**2. Results** High quality web surveys have response rates of about 3.5% (Pew Research Center, 2014). The response rate for our survey was almost twice that at 6.5% which, although small on an absolute level, is encouraging given our more challenging environment. In addition, unlike most web surveys, we were able to perform some checks for selection bias because we collected available information on our entire target sample, before administering our survey question. This information to test for selection bias included variables such as the number of followers, gender, year of creation, average number of posts for each month, and enabling geo-location; we also observed each of these variables within the five separate data sources in Section 4. The vast majority of tests we conducted indicated statistically insignificant differences between respondents and nonre-

spondents. The few differences that appeared are negligible compared to the large effect sizes we present below. As might be expected, the data contains some evidence that 50c party members are less likely to respond to our question than non-50c party members, which has the effect of making it more difficult to confirm our hypothesis. Although we could weight the results below by the differences we found, they are small enough that we chose to present the raw, unprocessed data below instead.

The results for our three surveys appear in Table 1. Overall, we found that 59% of our predicted 50c party members admitted to being 50c party members. If we are correct that they are all 50c party members, then the remaining 41% gave an insincere answer, which would not be surprising given that this is essentially their job. To test this, we use our sample of known 50c party members revealed in our leaked Zhanggong email archive. In this sample, 57% admitted to their 50c party status. The two percentage point difference between these two figures is not statistically significant (at  $\alpha = 0.05$ ), suggesting that indeed all respondents in our predicted sample are 50c.

Also as a test, we use our sample that approximates those known to not be 50c party members. In this sample, only 19% said they were 50c; the 40 percentage point difference between this figure and that from our predicted 50c party member sample (59%) is very large for a survey and statistically significant, revealing that we have detected a strong signal of actual 50c party membership among our predicted 50c sample. Overall, the results from this survey strongly support the validity of the predictions of 50c party membership conducted in Section 4.

| 50c Status      | Origin                   | Percent Yes |
|-----------------|--------------------------|-------------|
| Predicted 50c   | Across China             | 59%         |
| Known 50c       | Leaked Zhanggong Archive | 57%         |
| Known “Not” 50c | Random sample            | 19%         |

Table 1: Survey About 50c Status. The first line is our survey; the second two are gold standard evaluation surveys. The difference between the first and second lines is not statistically significant; the difference between the first and the third is statistically significant (both at  $\alpha = 0.05$ ).

## 6 Size of the 50c Party

In this section, we study how widespread 50c activity is. Overall, we find a massive government effort, where every year the 50c party writes approximately 448 million social media posts nationwide. About 52.7% of these posts appear on government sites. The remaining 212 million posts are inserted into the stream of approximately 80 billion total posts on commercial social media sites, all in real time. If these estimates are correct, a large proportion of government web site comments, and about one of every 178 social media posts on commercial sites, are fabricated by the government. The posts are not randomly distributed but, as we show in Figure 2, are highly focused and directed, all with specific intent and content. The rest of this section explains how we estimate these numbers. Throughout, in lieu of the possibility of formal standard error calculations, we offer transparent assumptions that others can easily adjust to check sensitivity or improve as more information is unearthed.

**1. Number of Social Media Posts** To understand the context into which 50c posts are inserted, we begin by estimating the total number of Chinese social media posts nationwide. As of December 2012, netizens were posting approximately 100 million messages a day, or 36.5 billion a year, on Sina Weibo alone (Zhao et al., 2014), which is one of at least 1,382 known social media sites (King, Pan and Roberts, 2013). In our data, the ratio of Sina Weibo posts to all posts is 1.85, meaning that an estimate of the total number of posts on all platforms is  $(1.85 \times 36.5 \text{ billion}) = 67.5 \text{ billion}$ . However, this requires the strong assumption that 50c party members use specific commercial social media platforms in the same proportions as the entire user population. We therefore use the detailed survey from iiMedia Research Group (2014) and calculate that the ratio of total posts to Sina Weibo posts to be 2.10, and the total number of posts per year to be about 80.4 billion. This is an underestimate because it is based on microblogs and ignores blogs, but blogs probably number in the millions which is rounding error on this scale.



**2. Number of 50c Posts in Zhanggong** Among the 43,797 confirmed 50c posts, 30,215 were made during a 365 day period between 2/11/2013 (the first day on which we observed a 50c post) and 2/10/2014. We have evidence of at least 1,031 exclusive (Sina Weibo) accounts in Zhanggong, including 202 accounts in the leaked archive and 829 we identified outside the archive (by following the rules in Section 4).

In our archive, a 50c party member needing to make a post chooses an exclusive account on Weibo ( $689/43,797 =$ ) 1.57% of the time compared to all other choices (an ordinary account on Weibo or another social media site). We assume this ratio is approximately the same for non-leaked 50c posts in Zhanggong, which in turn implies that the ratio of total 50c posts to 50c posts in the archive is the same as the ratio of total exclusive accounts to exclusive accounts in the archive. As such, an estimate of the total number of posts in Zhanggong in 2013 is  $(30,215 \times 1,031/202 =)$  154,216.

**3. Number of 50c Posts in Jiangxi Province** Zhanggong is an urban district of Ganzhou City within Jiangxi province. According to China Internet Network Information Center's Statistical Report on Internet Development in China 2014, the 2013 Internet penetration of urban residents was 62.0% and of rural residents was 27.5% (<http://j.mp/CNNIC>). According to the National Bureau of Statistics of China, 48.87% of the 45.22 million people in Jiangxi province lived in urban areas, or 22.10 million, with 23.12 million living in rural areas (<http://j.mp/ChinaSY>).

We first compute the number of 50c posts per internet user in Zhanggong, which is  $(154,216/468,461 \times 0.62 =)$  0.531. We then assume that this rate is roughly the same in Jiangxi and then scale up. Thus, we estimate a total number of 50c posts in Jiangxi during 2013 as  $(0.531 \times [0.62 \times 22.1M + 0.275 \times 23.1M] =)$  10.65 million.

**4. Number of 50c Posts in China** Finally, to scale this result to all of China, we assume that the ratio of 50c posts to population in other parts of China is roughly the same as in Jiangxi. This ratio of posts per user is  $(10.65M/14.68M =)$  0.7255. Applying this assumption to the country as a whole reveals the presence of  $(0.7255 \times 617.58M =)$  448.0 million 50c posts in China during 2013 (see <http://j.mp/CNNIC>).

## 7 Theoretical Implications

The empirical results offered above seem clear, but what do they suggest about the overall strategy of the Chinese government, or for authoritarian regimes in general? We first explain these results by generalizing prior findings on (human) censorship and (automated) filtering, all led by the same propaganda department in the same government as the 50c party (King, Pan and Roberts, 2013, 2014). We then extend these ideas to the authoritarian literature in general.

**1. China** One way to parsimoniously summarize existing empirical results about information control in China is with a theory of the strategy of the regime. This theory involves two simple and complementary principles the regime appears to follow, one passive and one active. The passive principle is *do not engage on controversial issues*, including no 50c posts supporting, and no censoring of posts criticizing, the regime, its leaders, or their policies.

To develop the active principle, we first define social media posts with “collective action potential” as those about real world crowd formation and related activities (see King, Pan and Roberts, 2013, p.6, for a more precise definition). Citizens criticize the regime without collective action on the ground in many ways, including even via (unsubstantiated) threats of protest and viral bursts of online-only activity — all which do not have collective action potential by this definition and so are ignored by the government. Then the second, active, principle is *stop discussions with collective action potential*, by active distraction and active censorship. Cheerleading in directed 50c bursts is one way the government distracts the public, although this activity can also be used to distract from other events, general negativity, specific grievances, etc.<sup>6</sup>

These principles appear to derive from the fact that the main threat perceived by the Chinese regime in the modern era is not military attacks from foreign enemies but rather uprisings from their own people. Staying in power involves managing their government

---

<sup>6</sup>Unfortunately, although our leaked archive includes directions to 50c workers, the information does not reveal whether these directions originate from Zhanggong or from higher levels of the government or party. This prevents us from determining with any certainty the specific intent of each burst of 50c activity.

and party agents in China's 32 provincial-level regions, 334 prefecture-level divisions, 2,862 county-level divisions, 41,034 township-level administrations, and 704,382 village-level subdivisions, and somehow keeping in check collective action organized by those outside of government. The balance of supportive and critical commentary on social media about specific issues, in specific jurisdictions, is useful to the government in judging the performance of (as well as keeping or replacing) local leaders and ameliorating other information problems faced by central authorities (Dimitrov, 2014a,b,c; Wintrobe, 1998). As such, avoiding any artificial change in that balance — such as from 50c posts or censorship — can be valuable.

Distraction is a clever and useful strategy in information control in that an argument in almost any human discussion is rarely an effective way to put an end to an opposing argument. Letting an argument die, or changing the subject, usually works much better than picking an argument and getting someone's back up (as new parents recognize fast). It may even be the case that the function of reasoning in human beings is fundamentally about winning arguments rather than resolving them by seeking truth (Mercier and Sperber, 2011). Distraction even had the advantage of reducing anger compared to ruminating on the same issue Denson, Moulds and Grisham (2012). Finally, since censorship alone seems to anger people Roberts (2015), the 50c astroturfing program has the additional advantage of enabling the government to actively control opinion without having to censor as much as they might otherwise.

Our inference about distraction being the goal of the regime is consistent with directions to 50c party members in emails from the Zhanggong propaganda department. They ask the 50c members to “promote unity and stability through positive publicity” (坚持“团结稳定 鼓劲、正面宣传为主”) and “actively guide public opinion through emergency events” (积极稳妥做好突发事件舆论引导). In this context, “emergency events” are events with collective action potential.<sup>7</sup>

---

<sup>7</sup>For example, a website developed by the Ministry of Public Security and Ministry of Education to help young people better understand safety issues (<http://j.mp/EmergEvents>) explains: “Every emergency event involves the self interest of a particular group of people, leading to psychological pressure and change among this group, and understandably leading to concern and worry. Especially for emergency events of a societal nature [as distinct from natural disasters], most are organized by a small group of people, who through their publicity-seeking and encouragement get others involved. Recently, emergency events

**2. Authoritarian Politics** For the literature on authoritarian politics in general, our results may help refine current theories of the role of information, and in particular what is known as “common knowledge,” in revolutionary mobilization. Many theories in comparative politics assume that autocrats slow the spread of information critical of the regime in order to minimize the development of common knowledge of grievances which, in turn, reduces the probability of mobilization against the regime. The idea is that coordination is essential to revolution, and coordination requires some common knowledge of shared grievances (Chwe, 2013; Egorov, Guriev and Sonin, 2009; Hollyer, Rosendorff and Vreeland, 2014; Persson and Tabellini, 2006; Tilly, 1978).

In contrast, our results seem to indicate that the Chinese regime differentiates between two types of common knowledge — about specific grievances, which they allow, and about posts with collective action potential, which they do a great deal to avoid. Avoiding the spread of common knowledge about collective action events (and not grievances) is consistent with research by Kuran (1989, 1991), Lohmann (1994), and Lorentzen (2010), who focus specifically on the spread of information about real-world protest and on-going collective action rather than the generic spread of common knowledge more broadly.

The idea is that numerous grievances of a population ruled autocratically by non-elected leaders are obvious and omnipresent. Learning of one more grievance, in and of itself, should have little impact on the power of a potential revolutionary to ignite protest. The issue, then, appears not to be whether such grievances are learned by large enough numbers to foment a revolution. Instead, we can think of creative political actors, including those aspiring to lead a revolution or coup, as treating issues, ideologies, events, arguments, ideas, and grievances as “hooks on which politicians hang their objectives and by which they further their interests,” including interests that entail initiating or fostering a political uprising (Shepsle, 1985). If one hook isn’t available, they can use another.

By this logic, then, common knowledge of grievances is already commonplace and so

---

due to issues like territorial disputes, land requisitions, and housing demolition in certain areas are often organized by one person and involve many, making collective events.” (任何一类突发事件，都必然要涉及一部分人的切身利益，使其产生心理压力和变化，引起人们的关注和不安也属正常。尤其是社会性的突发事件，多是由少数人操纵，通过宣传鼓动把一些群众卷到事件中来。近期，在一些地方因地界、征地、拆迁安置而发生的突发性事件，往往是一人纠合，数人响应，使其具有聚众性。.)

allowing more information about them to become public is of little risk to the regime or value to its opponents. Since disrupting discussion of grievances only limits information that is otherwise useful to the regime, the leaders have little reason to censor it, argue with it, or flood the net with opposing viewpoints. What *is* risky for the regime, and therefore vigorously opposed through large scale censorship and huge numbers of fabricated social media posts, is posts with collective action potential.

## **8 Concluding Remarks**

The vast majority of scholars, journalists, activists, and participants in social media have, until now, been convinced that the massive 50c party is devoted to engaging in argument that defends the regime, its leaders, and their policies. Our evidence indicates the opposite — that the 50c party engages in almost no argument of any kind and is instead devoted primarily to cheerleading for the state, symbols of the regime, or the revolutionary history of the Communist Party. We interpret these activities as the regime’s effort at strategic distraction from collective action, grievances, or general negativity, etc.

It also appears that the 50c party is mostly composed of government employees contributing part time outside their regular jobs, not, as has been claimed, ordinary citizens paid piecemeal for their work. This, nevertheless, is still an enormous workforce that, we estimate, produces 448 million 50c posts per year. Their effectiveness appears maximized by the effort we found of them concentrating the posts into spikes at appropriate times, and by directing about half of the posts to comments on government web sites.

Inferences in this paper depend on the veracity of the leaked archive we analyze. The size and extraordinary complexity of this archive makes it highly likely to be real. There are no signs of it having been generated by automated means, and doing so by hand would have been a monumental task. We have also verified numerous external references from the data — to specific individuals, email addresses, phone numbers, government departments, programs, web sites, social media accounts, specific posts, etc. — and every one checks out.

Of course, 50c party efforts may exist that do not follow the model we unearthed in

Zhanggong, and it is possible that such hypothetical organizations may follow different rules and practices, perhaps in different places, and of course may generate 50c posts with unrelated content. We think this is unlikely, but regardless, perhaps the window this paper opens on this enormous and previously secretive government program may help others discover different aspects of it in China, and eventually in other related authoritarian regimes.

## A Appendix: Categorization Scheme

Our categorization scheme for social media posts includes the six categories below, along with examples of each. Non-Chinese speakers should be aware when reading these examples that the Chinese language, even on social media, tends to be quite flowery and formal, with frequent creative, and often (to English speakers) stagy-sounding, wordings.

**(1) Taunting of Foreign Countries** Favorable comparisons of China to other countries; includes insults to other countries, and taunting of pro-democracy, pro-West, pro-individual liberties, or pro-capitalist opinions within China. Examples from leaked Zhanggong 50c posts:

- 去年，奥巴马在香格里拉会议上力邀23国参与围堵中国时这样说道：“中国有13亿人，他们越崛起，我们就会越没饭吃，因为地球资源供给是有上限的。所以为了我们能继续过现在的生活，就必须遏制中国的发展。” [Last year, at the Shangri-la Dialogue where Obama invited 23 countries to participate in the containment of China, he said: “China has 1.3 billion people, the faster China rises, the more difficult it will be for us to live, because the earth’s resources are limited. For us to remain at our current living standard, we must contain China’s development.”]
- 中国的崛起大势已经不可阻挡。美国一边公开宣称不是中国死就是西方亡，一边又拼命告诉中国民众：你们的政府有问题啊，必须推翻它，然后你们就能过上比现在更好的日子。——请问，还有比这更可笑和自相矛盾的逻辑吗？ [China’s rise is now inevitable. On one hand, the US publicly asserts that if China does not perish the West will wither; on the other hand it tells the Chinese people that your government is problematic: you have to overthrow it so you can live a better life. Is there a more ridiculous and contradictory logic than this?]

**(2) Argumentative praise or criticism** Comments on controversial, Pro/Con (non-valience) issues, as well as claims of wrongdoing or unfairness, praise (usually of the government) or criticism (usually of opponents of the government), taking a position or explaining why a particular viewpoint is correct or (more often) wrong. These posts are often part of a debate, in opposition to a previous post. Examples from leaked Zhanggong 50c posts:

- 我亲爱的朋友们，翻一下你的微薄，你会发现系统已经自动帮你添加了诸如薛蛮子、李开复、作业本、韩寒、李承鹏等各种民粹微薄，这是标准的强制灌输和洗脑手段，建议你取消关注 [My dear friends, you if you go through your Weibo, you'll discover that the system automatically had you follow Xue Manzi, Li Kaifu, Zuo Yeben, Han Han, Li Chengpeng and other populist Weibo users. This is a typical tactic of indoctrination and brainwashing, I suggest you unfollow them.]
- 李开复说纽约60万美金一套别墅，比北京便宜多了，但他不会告诉你那套所谓的别墅其实是个仓库，而且离纽约市区开车需要四个多小时 [Li Kaifu says that you can buy a villa for \$600,000 USD in New York, much cheaper than in Beijing. But what he doesn't tell you is that this so-called villa is actually a warehouse, which is more than a four hour drive from New York City.]

**(3) Non-argumentative Praise or Suggestions** Focused on noncontroversial valience issues which no one could argue against, such as improving housing and public welfare. It includes praise of current government officials, programs, or policies. It doesn't respond to alternative, opposing viewpoints, and it includes positive sentiment. It is distinguished from category (3) in that it praises something specific such as the government, its officials, government programs, or initiatives. It includes a small number of constructive suggestions for what government policies might include (i.e., added benefits rather than critical complaints). It does not argue against a specific viewpoint, but just says "it would be nice if the government did X," which usually the government is already in the process of implementing. Some examples of known Zhanggong 50c posts:

- 政府...做了好多实事，其中解决了好大一部分的人住房。 [The government has done a lot of practical things, among which is solving a significant part of the housing problem]
- 土坯房改造政策，让我们村的人搬出了坯房都住上了小洋楼，村里发生了翻天覆地的变化，真是太感谢了。 [The policy of renovating mud-brick houses has allowed villagers to move out of mud-brick dwellings into small, Western-style buildings. The village has been transformed, we are so grateful]
- 希望中央支持力度更大 [We hope the central government provides us with even more support]
- 希望能有更多类似《若干意见》的好政策！ [We hope there will be more good policies like "Various Opinions" (the abbreviated name of an economic development policy)]
- 期待书记带领我们...特别是在教育、医疗卫生方面，争取更多政策为百姓谋取更大福祉。 [We look forward to the leadership of our party secretary... We hope that he can carry out more policies that will benefit the people in different aspects, especially in education and health care.]

**(4) Factual Reporting** Descriptions of current government programs, projects, events, or initiatives, or planned or in progress initiatives. Does not include any praise of these programs or events (which would be category (3)), just that they are occurring. Reporting on what government, government officials are doing. Some examples of known Zhanggong 50c posts:

- 清明三天假期7座小客车继续免费 [During the Qingming festival three-day holiday, [the freeway] will remain free to 7-seater busses]
- 6月27日，江西省委作出向龚全珍同志学习的决定，号召全省党员干部深入学习龚全珍坚定信念、永葆本色的坚强党性，服务群众、一心为民的质朴情怀，忠于职守、矢志不渝的执着追求，淡泊名利、无私奉献的高尚情操，勤俭节约、艰苦朴素的生活作风。 [On June 27, the Jiangxi provincial committee promulgated an opinion to learn from comrade Zhen Gongquan, calling on all provincial party members and cadres to study Zhen Gongquan's firm conviction, staunch support of the Party's spirit, service to the masses, straightforward dedication to the people, devotion to duty, abiding dedication, indifference to fame and fortune, selfless dedication to moral character and hardwork.]
- 1月16日，江西省委常委、赣州市委书记史文清将通过中国赣州网与网民在线交流，倾听网民意见、建议和诉求。 [On January 16, Jiangxi Party Committee Member and Ganzhou City Party Secretary Shi Wenqing will communicate with netizens on the China Ganzhou Web, to hear comments, suggestions, and demands from netizens.]

**(5) Cheerleading for China** Patriotism, encouragement and motivation, inspirational quotes and slogans, inspirational quotes from government officials, thankfulness, gratefulness, inspiration or thankfulness for historical and aspirational figures or events, and cultural references and celebrations (e.g., describes traditions, actions, suggestions for the community). Does not include positive sentiment toward particular government leaders or specific policies, but does include positive sentiment or general praise toward life, historical figures, model citizens (e.g., Lei Feng; Gong Quanzhen, a model teacher; Guo Chuhui, a patriotic villager), or China in general. Some examples of known Zhanggong 50c posts:

- 众多革命先烈们的英勇奋斗，缔造了我们今天的幸福生活！向英雄致敬。 [Many revolutionary martyrs fought bravely to create the blessed life we have today! Respect these heroes.]
- 向所有为中华民族繁荣富强做出伟大贡献的先人们致敬！人民英雄永垂不朽 [Respect to all the people who have greatly contributed to the prosperity and success of the Chinese civilization! The heroes of the people are immortal]
- 接过父辈、祖辈血染的红旗，坚定不移地跟着党走！ [[I will] carry the red flag stained with the blood of our forefathers, and unswervingly follow the path of the CCP!]
- 我们自己要更加努力，不等不靠，主动上前。 [We all have to work harder, to rely on ourselves, and to take the initiative to move forward.]



- 爱我中华 [I love China]
- 大家的日子都过好了，中国梦就实现了！ [[If] everyone can live good lives, then the China Dream will be realized!]
- 赣州加油哦 [Way to go Ganzhou]

(6) **Other** Irrelevant posts that are entirely personal, commercial (such as ads), jokes, or empty posts that forward information not included. This category is removed and conditioned on in all analyses in the paper.

## References

- Bambauer, Derek E. 2013. "Censorship v3. 1." *Internet Computing, IEEE* 17(3):26–33.
- Bandurski, David. 2008. "China's guerrilla war for the web." *Far Eastern Economic Review* 171(6):41.
- Barr, Michael. 2012. "Nation branding as nation building: China's image campaign." *East Asia* 29(1):81–94.
- Bremmer, Ian. 2010. "Democracy in Cyberspace." *Foreign Affairs* 89(6):86–92.
- Chen, Jidong, Jennifer Pan and Yiqing Xu. 2016. "Sources of Authoritarian Responsiveness: A Field Experiment in China." *American Journal of Political Science* 60(2):383–400.
- Chwe, Michael Suk-Young. 2013. *Rational ritual: Culture, coordination, and common knowledge*. Princeton University Press.
- Deibert, Ronald and Rafal Rohozinski. 2010. "Liberation vs. control: The future of cyberspace." *Journal of Democracy* 21(4):43–57.
- Denson, Thomas F, Michelle L Moulds and Jessica R Grisham. 2012. "The effects of analytical rumination, reappraisal, and distraction on anger experience." *Behavior therapy* 43(2):355–364.
- Dimitrov, Martin K. 2014a. "Internal Government Assessments of the Quality of Governance in China." *Studies in Comparative International Development* pp. 1–23.
- Dimitrov, Martin K. 2014b. "Tracking Public Opinion Under Authoritarianism." *russian history* 41(3):329–353.
- Dimitrov, Martin K. 2014c. "What the Party Wanted to Know Citizen Complaints as a "Barometer of Public Opinion" in Communist Bulgaria." *East European Politics & Societies* 28(2):271–295.
- Editors. 2013. "Cat and mouse: How China makes sure its internet abides by the rules." *The Economist* April 6.
- Editors. 2016. "Fifty cents." *China Digital Times* . <http://j.mp/cdt50cent>.
- Egorov, Georgy, Sergei Guriev and Konstantin Sonin. 2009. "Why Resource-poor Dictators Allow Freer Media: A Theory and Evidence from Panel Data." *American Political Science Review* 103(4):645–668.
- Freedom House. 2009. "China and the Internet: An Uphill Fight for Freedom." *Harvard International Review* pp. 68–73.
- Greitens, Sheena Chestnut. 2013. "Authoritarianism Online: What Can We Learn from Internet Data in Nondemocracies?" *PS: Political Science & Politics* 46(02):262–270.

- Groves, Robert M, Floyd J Fowler Jr, Mick P Couper, James M Lepkowski, Eleanor Singer and Roger Tourangeau. 2011. *Survey methodology*. Vol. 561 John Wiley & Sons.
- Haley, Usha. 2010. "China's Fifty Cent Party for Internet Propaganda." *Huffington Post* October 4.
- Han, Rongbin. 2015. "Manufacturing Consent in Cyberspace: China's "Fifty-Cent Army"." *Journal of Current Chinese Affairs* 44(2):105–134.
- Hassid, Jonathan. 2012. "Safety valve or pressure cooker? Blogs in Chinese political life." *Journal of Communication* 62(2):212–230.
- Henochowicz, Anne. 2014. "Thousands of Local Internet Propaganda Emails Leaked." *China Digital Times* . <http://j.mp/leakedEmails>.
- Hollyer, James R., B. Peter Rosendorff and James R. Vreeland. 2014. "Measuring transparency." *Political analysis* 22(4):413–434.
- Hopkins, Daniel and Gary King. 2010. "Improving Anchoring Vignettes: Designing Surveys to Correct Interpersonal Incomparability." *Public Opinion Quarterly* pp. 1–22. <http://gking.harvard.edu/files/abs/implement-abs.shtml>.
- iiMedia Research Group. 2014. "2013 年中国移动社会化分享年度报告." *China Mobile Social Share Annual Report (内部版)* . <http://j.mp/iimedia>.
- Keating, Joshua. 2011. "China's '50-cent' party takes on the jasmine revolutions." *Foreign Policy March*.
- King, Gary, Jennifer Pan and Margaret E. Roberts. 2013. "How Censorship in China Allows Government Criticism but Silences Collective Expression." *American Political Science Review* 107:1–18. <http://j.mp/LdVXqN>.
- King, Gary, Jennifer Pan and Margaret E. Roberts. 2014. "Reverse-engineering censorship in China: Randomized experimentation and participant observation." *Science* 345(6199):1–10. <http://j.mp/1KbwkJJ>.
- King, Gary, Patrick Lam and Margaret Roberts. 2016. "Computer-Assisted Keyword and Document Set Discovery from Unstructured Text."
- Kuran, Timur. 1989. "Sparks and prairie fires: A theory of unanticipated political revolution." *Public Choice* 61(1):41–74.
- Kuran, Timur. 1991. "Now out of never: The element of surprise in the East European revolution of 1989." *World politics* 44(1):7–48.
- Lam, Oiwan. 2012. "China: Challenging the 50 Cent Party." *Global Voices March*.
- Lam, Oiwan. 2013. "China Beefs Up '50 Cent' Army of Paid Internet Propagandists." *Global Voices October*.
- Letham, Benjamin, Cynthia Rudin, Tyler H McCormick, David Madigan et al. 2015. "Interpretable classifiers using rules and Bayesian analysis: Building a better stroke prediction model." *The Annals of Applied Statistics* 9(3):1350–1371.
- Lohmann, Susanne. 1994. "The Dynamics of Informational Cascades: The Monday Demonstrations in Leipzig, East Germany, 1989-1991." *World Politics* 47(1):42–101.
- Lorentzen, Peter. 2010. "Regularizing Rioting: Permitting Protest in an Authoritarian Regime." Working Paper.
- Mercier, Hugo and Dan Sperber. 2011. "Why do humans reason? Arguments for an argumentative theory." *Behavioral and brain sciences* 34(02):57–74.
- Miller, Blake Andrew Phillip. 2016. "Automatic Detection of Comment Propaganda in Chinese Media." *Available at SSRN 2738325* .

- Näher, Anatol-Fiete and Ivar Krumpal. 2012. "Asking sensitive questions: the impact of forgiving wording and question context on social desirability bias." *Quality & Quantity* 46(5):1601–1616.
- Ng, Jason. 2011. "Wu Mao." *'Blocked on Weibo' blog* . blockedonweibo.tumblr.com.
- Persson, Torsten and Guido Tabellini. 2006. "Democracy and development: The devil in the details." *American Economic Review* 96:319–324.
- Pew Research Center. 2014. "Beyond Red vs. Blue: The Political Typology." <http://j.mp/PewResponse>.
- Philipp, Joshua. 2015. "Leaked Emails Show Chinese Regime Employs 500,000 Internet Trolls." *Epoch Times* . <http://j.mp/twomill>.
- Roberts, M. 2015. "Experiencing Censorship Emboldens Internet Users and Decreases Government Support in China." *Unpublished Working Paper*, URL <http://www.margaretroberts.net/wp-content/uploads/2015/07/fear.pdf> .
- Shepsle, Kenneth A. 1985. Comment of Why the Regulators Chose to Deregulate. In *Regulatory Policy and the Social Sciences*, ed. Roger Noll. Berkeley: University of California Press pp. 231–239.
- Sonnad, Nikhil. 2014. "Hacked emails reveal China's elaborate and absurd internet propaganda machine." *Quartz* . <http://j.mp/Sonnad>.
- Strafella, Giorgio and Daria Berg. 2015. "'Twitter Bodhisattva': Ai Weiwei's Media Politics." *Asian Studies Review* 39(1):138–157.
- Tang, Min, Laia Jorba and Michael J Jensen. 2012. "Digital media and political attitudes in China." *Digital Media and Political Engagement Worldwide: A Comparative Study* pp. 221–239.
- Tilly, Charles. 1978. *From mobilization to revolution*. McGraw-Hill New York.
- Tong, Yanqi and Shaohua Lei. 2013. "War of position and microblogging in China." *Journal of Contemporary China* 22(80):292–311.
- Tourangeau, Roger, Frederick Conrad and Mick Couper. 2013. *The science of web surveys*. Oxford University Press.
- Wintrobe, Ronald. 1998. *The Political Economy of Dictatorship*. Cambridge University Press.
- Yang, Xiaofeng, Qian Yang and Christo Wilson. 2015. Penny for Your Thoughts: Searching for the 50 Cent Party on Sina Weibo. In *Ninth International AAAI Conference on Web and Social Media*.
- Zhang, Qingpeng, Dominic DiFranzo and James A Hendler. 2014. Social Networking on the World Wide Web. In *Encyclopedia of Social Network Analysis and Mining*. Springer pp. 1879–1892.
- Zhao, Juanjuan, Weili Wu, Xiaolong Zhang, Yan Qiang, Tao Liu and Lidong Wu. 2014. "A short-term trend prediction model of topic over Sina Weibo dataset." *Journal of Combinatorial Optimization* 28(3):613–625.