Article:
Bolsover, G orcid.org/0000-0003-2982-1032 (2018) Slacktivist USA and Authoritarian China? Comparing Two Political Public Spheres with a Random Sample of Social Media Users. Policy and Internet, 10 (4). pp. 454-482. ISSN 1944-2866

https://doi.org/10.1002/poi3.186

Slactivist USA and Authoritarian China? Comparing Two Political Public Spheres with a Random Sample of Social Media Users

Gillian Bolsover [gillianbolsover@gmail.com], School of Politics and International Studies, University of Leeds, Woodhouse Lane, Leeds, LS2 9JT.

The rise of social media has put back on the agenda questions about the Internet's potential as an online public sphere, particularly in authoritarian states. However, random samples have never been employed to investigate political speech on social media, necessarily limiting knowledge. This article presents an analysis of political speech based on a random sample of more than 1,000 active US Twitter users and Chinese Weibo users collected in late 2014. Political speech by ordinary users was found to be more frequent on both platforms than expected – 9.4% on Weibo and 6.8% on Twitter - lending support to hopes of an online public sphere. However, existing powerholders make up around a fifth of US Twitter accounts, and political speech acts by ordinary US Twitter users are largely "slacktivist" in nature. In contrast, 98% of active accounts on Weibo belong to ordinary users, with active political speech making up more than one in fifty posts by these users. Although they largely fall within the bounds of what is permitted by the Chinese state, these findings point to the potential of the Internet as a (limited) public sphere in China, while raising questions about its contribution to political processes in the US.

KEYWORDS: China, social media, Twitter, Weibo, public sphere, US, political speech, random sample, elite domination

The Internet and Politics in Democratic and Authoritarian States

The Internet, when it first became popular in the 1990s, was seen as having the potential to provide a space for the development of an online public sphere, reinvigorating offline civic life that was languishing with individualization, entertainmentization, and commercialization (Poster, 1997; Rheingold, 1993). By transferring power to networked individuals, it was thought that the Internet could incite a "control revolution," bypassing gatekeepers and allowing access to more diverse information (Shapiro, 2000). By enabling increased access to information, and lowering barriers to communication and connection with like-minded individuals, the Internet was seen as having the potential to reinvigorate established democracies, as well as to undermine authoritarian states, including China. Several studies have posited the emergence of an online civil society and public sphere in China—albeit largely within state-set boundaries (Yang, 2009; Yang and Calhoun, 2007; Zheng and Wu, 2005; Zhou et al., 2008).

Discussion about the liberation potential of the Internet in China and other authoritarian contexts persisted even after the hopes of the Internet's potential in democratic contexts started to sour (Diamond, 2010). Far from being democratizing, it was found to be increasingly controlled by existing power structures and dominated by a small number of offline power holders and a new online elite. An influential study found that only 0.12 percent of Internet traffic went to political websites (Hindman, 2009). The rise of social media, whose business model rests on the monetization of user data and online socializing, coupled with the concentration of power in the hands of a few service providers (Lovink, 2011), only served to exacerbate this trend. While these "Web 2.0" platforms provide affordances for political speech, their participatory potential is undermined by colonization by the market; censorship by organizations, states and industries; and appropriation by political and cultural elites (Cammaerts, 2008).

However, this dystopian narrative was soon challenged by the "Arab Spring"—the wave of protests that swept the Middle East and North Africa in 2011. Platforms that had been understood to lead to "echo chambers," "slacktivism," surveillance, distraction, and exploitation were now seen as playing a crucial role in facilitating political action: shaping the contours of political debate, facilitating organization of offline activities, and helping spread political ideas across borders (Howard et al., 2011; Howard and Hussain, 2011). Subsequent events—including the Black Lives Matter movement (Carney, 2016), and recent anti-government protests in Zimbabwe (Young, 2016)—have put back on the agenda questions about whether the Internet, and in particular the social media platforms that dominate today's online activities, might indeed empower individual users to participate in political speech. An additional question rased by these events is the extent to which

any effects on political speech might differ between democratic and non-democratic contexts.

Existing Research and its Limitations

Despite much attention paid to the Internet and politics by the academic community and policy makers, there are two key limitations to existing knowledge. Firstly, hardly any of the research on politics and social media is based on random samples. Due to the difficulty of constructing and collecting random samples, and the size of the datasets generated, the few studies based on them tend to be mostly descriptive, and rely on data generated by the platform rather than examining who uses social media and why. These studies include Fu and Chau's (2013) work on Weibo, and Gerlitz and Rieder's (2013) work on Twitter—neither of which address questions related to politics.

Rather than drawing on random samples, essentially what we know thus far about social media's role in political events is based on trending topics, or particular key words, users, groups, or case studies. This body of work has demonstrated Twitter's importance in a number of political events from the Green Revolution in Iran (Burns and Eltham, 2009), to the Indignados in Spain (González-Bailón et al., 2011), the Occupy movement in the US (Penney and Dadas, 2013), and the 2011 Egyptian protests (Howard et al., 2011). Research has also shown that Sina Weibo holds (or held) a similar position in China, due to its ease of use, massive user base, and rapid dissemination of information (Poell et al., 2013; Sullivan, 2014). A number of case studies, including the "Grass Mud Horse" meme (Tang and Yang, 2011), a suspicious fire near the state TV's headquarters (Sullivan, 2012), and the Wenzhou train crash (Nip and Fu, 2016), have indicated the importance social media platforms as venues for political speech in China.[1]

However, researchers have raised concerns that a reliance on streaming APIs to collect data introduces bias (Driscoll and Walker, 2014; González-Bailón et al., 2014; Morstatter et al., 2014), and it has been shown that datasets that rely on hashtags to collect posts on a specific topic produce substantively different results when compared with datasets that collect the post streams of users (D'heer et al., 2017). Thus, in order to better understand the role of social media as a potential public sphere, we need a dataset based on the user streams of randomly selected accounts.[2] This would allow us to assess how social media might be used for

[1] (Sina) Weibo has been so important in China that the Internet is seen by some as divided into two eras: pre-Weibo and Weibo. Although there has been much hope and many studies of the interaction between Weibo and political speech, many of the most iconic cases of Internet activism in China actually occurred in the pre-Weibo era (Yang, 2015, p. 3).

[2] Previous studies of trending topics have attempted to provide a picture of the kinds of speech in which users of microblogs engage. Based on an analysis of trending topics, Kwak et al. concluded that Twitter is a news medium rather than a social network (2010). Replicating this work on Weibo, Asur et al. concluded that the platform was not a news

political functions in a general sense, as opposed to in specific cases, where hashtag or keyword collection might indeed more closely resemble the underlying conversation about a given topic.

A second key limitation of research is the way in which different national and political contexts are approached; most of the research in communications and Internet studies focuses on Western, democratic contexts (Gunaratne, 2010; Waisbord and Mellado, 2014). A significant amount of research does address the interaction between the Internet and politics in non-Western and non-democratic contexts; the political effects of the Internet in China is a particularly large research area. However, these research efforts often either apply democratic, Western-generated theories unquestioningly in these new contexts or throw out existing theories based on a presumption of exceptionalism (Bolsover, 2017b). It is imprudent to apply research based on studies of Western, democratic countries in other contexts without questioning its applicability, but neither is it rational to believe that it will be completely irrelevant. Comparative research is the best way to understand the effects of political context on how the Internet might be used for political functions (Kennedy, 2011).

There is a current contradiction in how online platforms and speech acts are understood in different political systems. The same platforms and actions that are labelled as "slacktivist" and detracting from more desirable forms of political speech in democratic contexts (Morozov, 2013) are often seen as potentially highly positive for political speech in authoritarian contexts. For instance, memes and political humor are generally dismissed as "slacktivist" in democracies, but are hailed as a positive contribution to individual political powers in authoritarian countries, such as Vietnam (Sharbaugh and Nguyen, 2014) and China (Tang and Yang, 2011). It is impossible, however, to understand how social media platforms, in a general sense, are being used for political speech without the use of direct comparisons in different kinds of political systems. Thus, this research will approach these questions about whether and how social media acts as a venue for political speech based on a comparison of two random samples of Internet users and their speech acts on online platforms in the US and China.

The US and China are chosen as cases of comparison because, while they sit on opposite poles of the political spectrum, they are among only a handful of countries in which the national political, economic, and social contexts can be seen reflected in the online platforms used by their citizens. Across most of the world, platforms originally created in the US and reflecting these conditions of creation

---

medium but was instead dominated by frivolous content (2011). However, attempting to understand the underlying conversation patterns of microblog users based on trending topics encounters the same problems as hashtag tracking. Additionally, the way in which trending topics actually work is opaque and does not match the ideas that users have about this being a representation of a free market for information (Bolsover, 2017a).

predominate. China, Russia, and Iran are among the very small number of countries that have significant alternative national online spheres. Thus, the US and China are selected as countries of comparison because of their opposing political systems and the similarity of their online spaces in reflecting their political, economic, and social contexts of use.

Within each of these two countries, two social networking platforms are selected for comparison: Twitter in the US and Sina Weibo in China. Both Twitter and Weibo are microblogging platforms, which allow users to publish short messages that are disseminated to their network of followers and that other users can choose to repost or comment on.[3] Due to these affordances, both Twitter and Weibo are seen as particularly important venues for online political speech in their respective countries, as the wealth of literature addressing these platforms has shown. Although the crackdown on online rumors in China that started in late 2012 seems to have undermined some of the political functions of microblogs and encouraged individuals to move toward private services (Moore, 2014), it is still the case that microblogs, and Weibo in particular, are (one of) the main potential venues for public speech in the country.

Research Questions and Hypotheses

The introduction has highlighted some persistent questions about whether social media might function as a space for political speech, and whether they might also assist in the development of an online public sphere in authoritarian countries; in particular China. Some of the main critiques of the capacity of social media to act as a public sphere have been the influence of existing powerholders in society (state and market entities, and political and cultural elites), low use of these spaces for political functions, and the undermining of political speech by "slacktivism."

One reason why these questions persist is the lack of studies drawing on random samples, and a lack of comparative research. This article therefore addresses the question of the ability of social media to act as an online public sphere based on a random sample of Twitter users in the US and Weibo users in China. This is broken down into five research questions (RQ 1–5), discussed below:

RQ 1: To what extent are microblogs dominated by existing powerholders?

---

[3] Indeed, Weibo was launched in 2009 (three years after Twitter) as a Twitter clone. However, its functionality has now grown to exceed that of Twitter, with greater multimedia integration, threaded commenting systems, and an internal wallet (among other additional functionalities). Although neither of these sites is the largest website nor the largest social networking site within the country, they are both the leading one-to-many communication platforms in their national contexts. Other popular platforms such as Facebook, Instagram, QQ, and WeChat are closed; they may be used by a larger percentage of the population but their private nature means that it is difficult to see them as a potential public sphere.

RQ 2: Is there a difference in the extent of domination by existing powerholders between Twitter in the US and Weibo in China?

Colonization by the market and appropriation by political and cultural elites are two reasons why social media are failing to live up to their participatory potential (Cammaerts, 2008). If social media are dominated by existing elites—from the state, market, or civil society—then these platforms cannot act as an online public sphere (Habermas, 1989). Indeed, early research into online representation (pre-social media) concluded that rather than radically changing the balance of power in society, the Internet simply transfers offline power hierarchies online (Hindman, 2009). That said, the dearth of random samples means we do not have an accurate picture of user representation on microblogs. Although some studies have assessed user representation based on quantitative data from randomly selected social media profiles, such as verified status, gender, and location on Weibo (Fu and Chau, 2013), these efforts can't speak to the types of voices represented on the platform.[4]

Despite this lack of research, there is reason to believe that the extent of domination by existing elites may differ between political systems. In the US, Twitter markets itself as a medium for individual expression whereas, in China, Weibo actively courts celebrities and bloggers with its Media Microblogging contracts (Bolsover, 2017a). A number of studies have pointed to a leading role for opinion leaders as information intermediaries on Chinese social media (Bolsover, 2013; Fu and Chau, 2014; Liu, 2011), and the Chinese Communist Party has also promoted the use of Weibo to communicate directly with citizens (Song, 2015). Thus, we hypothesize that both Twitter and Weibo would likely be dominated by existing elites, but that Weibo might have a larger population of celebrities, bloggers, news media, state outlets, and opinion leaders (RQ 1-2).

RQ 3: How much of what microblog users post online can be characterized as political?

The prevailing "myth of digital democracy" suggests that only a very small fraction of online activity is political (Hindman, 2009). One study that coded a random selection of 200 posts from the Twitter "garden hose" found no instances

[4] Offline surveys often provide the best available data on user representation and speech on microblogs. For instance, in the US, the Pew Research Center publishes regular survey data, generally collected using random digit dialling, that provide a picture of the demographic profile of individual microblog users (Duggan and Brenner, 2013; Greenwood et al., 2016). However, this doesn't speak to the representation of different voices on social media, and comparable data are not available for many other national contexts.

of topical information about government and politics (Ghosh et al., 2013). However, the study's small sample size, and lack of a specified coding frame and trained coders, means that, although there is evidence of a general lack of political content on microblogs, there is a clear need for more research that specifically addresses this question (RQ 3).

RQ 4: How can the way that individuals use microblogs for political speech be characterized?

RQ 5: Is there any difference between rates or types of political speech between Twitter users in the US and Weibo users in China?

Political activities on social media are often criticized as "slacktivist," making "online activists feel useful and important while having preciously little political impact" (Morozov, 2012, 190). However, apart from the study by Ghosh et al. (2012) noted above, no research has analyzed the nature or frequency of political postings on microblogs based on a random sample. Surveys in the US have found that 66 percent of social media users report using these platforms for political activities, and 31 percent report having used them to encourage other people to take action on a political or social issue (Rainie et al., 2012). While these data indicate a high level of political activity, they only record the status of ever having participated in an online political activity and, thus, only contribute a small amount of information about the landscape of discourse on social media. These data might also be subject to a desirability bias, leading individuals to over-report political actions.

Interestingly, cross-national survey data have suggested that users in emerging nations participate in online political speech more frequently than those in established democracies[5]—users in China report expressing political opinions online no less frequently than other emerging nations, despite higher levels of censorship in the country (Bolsover et al., 2014). It has previously been hypothesized that this might be because of, rather than despite, the authoritarian nature of the Chinese state: "in democratic countries, besides the Internet people have other channels to express their opinions and to participate in politics, whereas in China the Internet is perhaps the single most important avenue for people to criticize government policies and to participate in politics" (Zheng and Wu, 2005, 525). Thus, we hypothesize that while the majority of online political speech may indeed be "slacktivist," it may be the case that rates of political speech are higher

[5] These cross-national surveys might be more helpful for hypotheses generation in relation to differences in rates of political speech if the extent of desirability bias was assumed to be constant across countries.

on Weibo in China than on Twitter in the US, because the platform provides a better space for political speech than the spaces available offline (RQ 4-5).

Constructing a Random Sample of US Twitter Users and Chinese Weibo Users

This study aims to provide data about the representation and activity of regular, individual microblog users in the US and China, based on a random sample of monthly-active, non-private US-based Twitter users and monthly-active, non-private mainland China-based Weibo users.[6]

All Weibo users are associated with a 10-digit ID number, and Twitter users a 64-bit ID number. Custom Python scripts were used to query the APIs of Twitter and Weibo based on sets of randomly generated potential ID numbers. Based on estimates of the underlying population and the desired confidence level and confidence interval of the results, a target of at least 500 monthly-active, non-private individual user accounts was set for each platform.[7] Between July and September 2014, 95,148 ID numbers were queried on Weibo, and 269,837 ID numbers were queried on Twitter. Data collection was stopped once it became clear, based on analysis of previous batches of randomly selected ID numbers, that the target sample of 500 regular, individual users on each platform would be reached.

The vast majority of ID numbers queried did not correspond to user profiles. Of the 95,148 ID numbers queried on Weibo, 15,972 (16.8 percent) were associated with user profiles, of which 1,057 (1.1 percent) had posted in the last month, had listed their location as within mainland China, and were not set to private. Of the 269,837 ID numbers queried on Twitter, 105,721 (39.2 percent) were associated with user profiles, of which 10,073 (3.73 percent) had public profiles and had posted in the last month.

---

[6] We excluded the following accounts from the analysis: private accounts, given that they do not form part of an online public sphere; accounts that post less than once a month (i.e. below the user-base threshold as defined by both Twitter and Weibo), given that we seek to investigate user representation and speech acts in online discourse among ordinary users; and non-mainland Chinese Weibo accounts, given that users based in Hong Kong, Macau, Taiwan, and overseas, live under a different political systems.

[7] The total population of monthly-active, non-private US-based Twitter users and monthly-active, non-private mainland China-based Weibo users is ultimately unknown. Although both companies make frequent reports as to the size of their user base, it is often hypothesized that these numbers are inflated in order to attract profit. It was revealed in October 2017 that Twitter had been overstating its monthly-active user base for the past three years (Tracy, 2017). Reported numbers of monthly-active Weibo users are also often seen as inflated (Custer, 2011; Fu and Chau, 2013). However, these figures can be used to establish an upper bound on the size of the underlying population. For populations of size one million or larger, a sample of 384 is necessary to state results with a confidence level of 95 percent and a confidence interval of ±5 . (Krejcie and Morgan, 1970). A sample size larger than 384 enables higher confidence levels and lower confidence intervals; based on this, a target of at least 500 individual users was set for each platform.

Geolocation is trivial on Weibo,[8] as users must set their location to the city or county level during profile creation; these data are returned via the API. However, geolocation is extremely difficult on Twitter, as the use of geolocation services is low and there is no established mechanism for geolocating users (Graham, Hale, & Gaffney, 2013). Previous research using qualitative methodologies has succeeded in geolocating around 63–66 percent of samples; however, the number of users who input location information appears to have fallen since these studies were conducted (Hecht et al., 2011; Takhteyev et al., 2012).[9]

A three-stage process was constructed in which user-input locations, geolocated tweets, and user descriptions were considered in order to attempt to locate the country of origin of the randomly selected Twitter accounts. Of the 10,073 randomly selected, monthly-active, non-private Twitter accounts, 4,192 (41.6 percent) could be located to the US based on this three-stage process.

The strength of this sample would, of course, be greater if 100 percent of the profiles could be geolocated. Rates of geolocation lower than 100 percent have the potential to introduce bias into the underlying sample, as the population of users that can be geolocated may differ substantially from the population that cannot be geolocated. In order to assess the magnitude of this potential bias in this dataset, the differences between the sample of geolocated profiles and the original random sample was compared on available metrics.

There was a statistically significant difference between the average number of statuses posted in the geolocated subset and the initial sample, with the geolocated subset having posted more statuses. However, there was no difference in the number of friends or number of followers. A comparison of user time zones and platform languages suggested that users in Japan and the Middle East were underrepresented in the geolocated subset (but as this article is concerned only with US users this would not affect results). The proportion of users using American English as a platform language and the proportion of users in the three major US time zones is roughly the same in the initial random sample and the geolocated subset. This suggests that, at least with relation to US users and measuring from the

---

[8] Assuming, of course, that one trusts user-input location data.

[9] The difficulty of geolocation on Twitter has spawned a whole area of research that attempts to infer the location of users based on their network connections to users who can be geolocated, with degrees of success ranging from 11 percent to 100 percent depending on the methods used, the degree of preciseness of the location, and the desired certainty of the inferred user locations (Backstrom et al., 2010; Compton et al., 2014; Davis Jr et al., 2011; Jurgens, 2013; Kong et al., 2014; Li et al., 2012; Rout et al., 2013). This level of computational complexity is outside the bounds of this article because it would require the collection of a large network of friends and followers centered around each of the 10,073 randomly selected Twitter users. Additionally, there are significant ethical questions that surround the inference of randomly selected Twitter users' locations based on information provided by the accounts to which they are connected. Twitter users who choose not to make public information that identifies their location have a reasonable right to expect privacy with respect to their location.

available data, the proportion of US-based users in the geolocated subset and initial random sample are roughly similar. (These data are presented in more detail in Supplementary Online Data Table 1.)

Despite the limitations of manual geolocation, this sample still represents the only attempt to provide data about the representation of different kinds of voices among active, US-based Twitter accounts and the types of speech in which individual users engage. Profiles from 124 countries were identified in this random sample, of which the largest fraction (22.95 percent, 962 accounts) belonged to users located in the US. (Data on the worldwide spread of active Twitter accounts are provided in Supplementary Online Data Table 2.) Thus, after geolocation, this dataset consisted of 1,057 randomly selected, monthly-active, non-private, mainland China-based Weibo accounts, and 962 randomly selected, monthly-active, non-private, US-based Twitter accounts.

In the second stage of this data collection process, the online activity (posts and forwards) of these users was collected using custom Python scripts over a four-week period, to construct a dataset of the online speech acts of these randomly selected users on both platforms. The period of 8 October to 4 November, 2014, was chosen because this research is specifically interested in the use of these online platforms for political speech, and important (approximately) biennial political events occurred in both countries at the end of this period: the 2014 US midterm election was held on 4 November (the final day of this four-week period) and the Fourth Plenary Session of the 18th National Congress of the Communist Party of China occurred between October 20 and 23. This period was, thus, designed to capture online political speech about the formal political process that might not occur at other times due to the relative infrequency of these events.

Accounts that did not post during this data collection period as well as those that were deleted, changed their status to private, were unavailable for part of the data collection, or, on Weibo, unfollowed the research account, were not included in the final dataset. At the end of data collection, the full dataset consisted of 686 randomly selected US-based Twitter users and 580 randomly selected mainland China-based Weibo users and their posts over a four-week period.[10]

This dataset is unique in that it is the first attempt to provide a general picture of the types of speech acts engaged in by individual users on these key commercial social networking sites. The vast majority of efforts to investigate political speech on both Twitter and Weibo focus on specific cases, trending topics,

---

[10] In keeping with common practice in quantitative Internet research based on large, digitally collected datasets, no contact was made with individual users. Data presented in this article is done in an aggregate format that does not risk identification. On both platforms, research accounts, clearly labelled as such, were created that followed the randomly selected user accounts. On Weibo, users can choose to remove an account that has followed them and those accounts that removed the research account at any point before or during data collection are not included in the dataset.

or opinion leaders; however, this preselects for successful cases and, thus, analysis of a large dataset without preselecting for outcome is necessary to truly understand the nature of political speech on microblogs (Sullivan, 2014).

Elite Dominance? Twitter Shows Much Greater Representation of Existing Powerholders

After the construction of this random sample of user accounts and their speech acts, a content analysis was conducted to ascertain what type of user these accounts belonged to, in order to assess the extent to which representation on these platforms is dominated by existing powerholders. The accounts were coded as one of 13 types: individual; public individual; celebrity; business (more than 50 employees); small business (less than 50 employees); small group or individual content or product producers; media outlet; blog, forum or online directory; civic group; non-profit organization, charity, or professional or advocacy organization; university, school, or official university or school organization; government department or publically run entity; and robot or spam account.

The initial list of account types was based on previous research (Bolsover, 2013) but was updated during the coding process to ensure that the categories used were applicable and comparable across both the Twitter and Weibo datasets. In order to arrive at categories that properly described the dataset but also were rigorous and replicable, a subset of accounts in both datasets were first roughly categorized. Based on this, a coding scheme was created that could distinguish between different kinds of accounts, and the data re-coded according to this new scheme.

The categorizations for account type were verified by a second coder. The second coder coded a random selection of around a third of the Twitter accounts (239 accounts). The percentage agreement was 82 percent with a Krippendorf's Alpha of 0.60, which is within acceptable levels for this type of research (Lombard et al., 2002). On Weibo, due to a very high prevalence of individual accounts, the second coder coded a non-random selection of around a third of the dataset (214 accounts). All of the accounts that had been coded as not belonging to individuals were included in this set plus a random selection of 203 of the accounts that had been coded as individuals. The percentage agreement between the first and second coder for the Weibo dataset was 82 percent[11].

[11] The subset of accounts double-coded in the Weibo sample was not random. The percentage agreement of the 203 accounts coded as individuals by the first coder was 90 percent. Based on this, the estimated percentage agreement of the two coders across the whole Weibo dataset would have been 87 percent. Measures that take into account predicted percentage agreement are usually a better indication of coding validity; however, in the case of the Weibo dataset, where more than 90 percent of accounts belong to individuals, these measures are poor indications of agreement and

The results of this content analysis are shown in Table 1. A significantly higher proportion of accounts on Twitter belong to existing powerholders: 11 percent of non-robot accounts belong to a market-sphere entity and seven percent to a civic entity. In comparison, less than one percent of active, mainland Chinese Weibo accounts belong to either a market-sphere or civic entity, with 99 percent of accounts belonging to single users ($p<0.0001$).

Given the stronger civil society presence in the US, one might expect that a greater proportion of accounts on US Twitter would belong to civic groups. However, the dearth of market-sphere accounts on Weibo is relatively surprising, particularly given that many of the market-sphere accounts on Twitter belong to small businesses or individual business people. A large proportion of Chinese individuals are self-employed and the Internet is often framed in China in terms of its ability to spur economic development. However, businesses and individual business people do not appear to be using Weibo as a tool for promotion and communication in the same way as Twitter has been used in the US. It is possible that this demographic in China uses other online platforms, such as Taobao or Douban, or that, given more recent Internet adoption and lower penetration in China, these businesses have yet to move online.

Table 1. User account types

	Weibo		**Twitter**		**Significance of difference (two-tailed Fischer's Exact Test)**
	Number	Percentage of non-robot accounts (%)	Number	Percentage of non-robot accounts (%)	
Single user accounts					**<0.0001**
Individual	569	98.10	552	80.50	<0.0001

percentage agreement is an appropriate measure of intercoder reliability, which falls within acceptable bounds for this type of research.

Public individual	3	0.52	7	1.18	0.3590
Celebrity	2	0.34	2	0.34	1.0000
Market-sphere accounts					**<0.0001**
Business (more than 50 employees)	2	0.34	21	3.54	0.0002
Small business (up to 50 employees)	0	0	38	6.41	<0.0001
Individual business people or small groups of individuals	0	0	15	2.53	<0.0001
Civil society-sphere accounts					**<0.0001**
Media outlet	0	0	2	0.34	0.5031
Blog, forum or online directory	3	0.52	16	2.70	0.0093
Civic group	0	0	19	3.20	<0.0001
Non-profit, charity or professional advocacy organization	0	0	7	1.18	0.0177
University, school or official university of school organization	0	0	4	0.67	0.1297
State-sphere accounts					**0.6297**
Government department	1	0.17	3	0.51	0.6297

or publically run entity					
Total Accounts	580		686		

It was hypothesized, based on previous research, that Weibo would have larger populations of celebrities, bloggers, news media, state outlets, and opinion leaders. Thus, it is surprising that there were more than twice as many public individuals and three times as many state-sphere accounts on Twitter than Weibo. However, the small number of accounts of this type in the sample means that the differences in the number of public individuals and state-sphere accounts between the platforms are not statistically significant and, thus, no conclusions can be drawn from these differences.

In reference to research questions one and two, it can be concluded based on these data that Twitter has a greater representation of existing power holders in society, while Weibo is used almost exclusively by individuals.[12] It may be the case that the differences found are a function of the different trajectories of development of the sites, with Twitter launched in the US three years before Weibo was launched in China. However, it may also be the case, in keeping with hypotheses based on previous research, that Weibo is seen as offering a freer space for information and expression than Chinese offline spaces, which might attract individual users to the platform. For this reason, the remaining three research questions are concerned with how individuals use the platform.

Individual Speech? Twitter is More of a Social Network, Weibo is More Commonly Used as a Medium for Active Political Speech

The quantitative statistics associated with individual users' posts on the two platforms are very different. On Twitter, the average number of statuses posted per week by an individual user was 20, while on Weibo the average was less than five. A much larger number of posts by individuals on Weibo were forwards (57 percent compared to 39 percent on Twitter). This suggests that Twitter is used more frequently for individual speech acts but Weibo more for information

[12] The hypothesis that these differences might simply have been an artefact of the low level of geolocation on Twitter is based on the idea that market, state and civil society might be geolocated at a much higher rate than individual accounts. However, in a scenario in which 100 percent of US-based market, state and civil society accounts in the initial random sample were geolocated but only 41.6 percent of US-based individual accounts were, the percentages would (of course) change but the conclusions drawn from this comparison would be unchanged. In this (worse-case) scenario, it would remain the case that there were significantly more individual and single-user accounts on Weibo ($p<0.0001$) and significantly more market and civil society accounts on Twitter ($p<0.0001$ in both cases).

dissemination. The voices of verified users appear to be far more prominent on Weibo, with 29 percent of all posts by individual users being forwards of statuses originally posted by verified users. This figure is, however, not very surprising given that the percentage of accounts that are verified is much higher on Weibo; 2.6 percent of profiles in the random sample on Weibo belonged to verified users, compared to none of the profiles in the (larger) Twitter sample. This suggests that these smaller numbers of verified users on Twitter, who are often also prominent offline, may have much more influence in setting the online agenda. (Quantitative data related to individual user's online activities are provided in Online Supplementary Data Table 3).

These quantitative data speak to the frequency of individual speech acts and the dominance of these speech acts by existing power holders. However, they cannot distinguish the frequency of different types of speech act by individual platform users. In order to address this question, a random selection of 500 posts by individual users on both platforms during the four-week period under consideration was coded to ascertain the topical content of individual user's posts.

These 1,000 posts were selected randomly from within the sample of 43,543 posts by individual users on Twitter and 8,907 posts by individual users on Weibo.[13] Each of the posts was coded to ascertain the type of information being shared: personal (sharing information about the self, or personal messages to another user), commercial (entertainment, products and services, or employment), informational (inspirational or helpful content; memetic or viral content; or links to offsite, non-political information), political (commenting on society, sharing news, commenting on the political process, attempting to exert influence, or political humor), or spam. (Example posts for each of these categories are provided in Supplementary Online Data Table 4.)

Several iterations of the coding scheme were tested. The original coding scheme was crafted to include topics identified in existing research into types of online speech (e.g. Asur et al., 2011; Ghosh et al., 2013; Kwak et al., 2010). In this article, however, we are particularly interested in political speech and distinguishing between different types of political speech. Drawing from previous research by the author that formulated a context-neutral definition of political speech (Bolsover, 2017a), this research defines political speech as any communicative action that affects or seeks to affect the balance of power in society.

Within the category of political speech, the initial coding frame incorporated ideas about the three modes of politics that the Internet is seen as affecting—namely, information, mobilization and interaction (Lilleker and Thierry, 2013)—and key debates about how to define political participation in the modern

---

w

world (Fox, 2013). During a series of pre-tests on different random samples of the dataset, several categories (such as political humor) were added and others merged to more accurately reflect the actual nature of online speech on both platforms. The category of identity assertion as a form of political speech was dropped from the coding scheme because it was deemed by both coders to be too difficult to ascertain for an individual post without knowledge of the offline context of the speaker.

A second coder coded a random selection of 100 of the 500 tweets on each platform. The percentage agreement for the five broad categorizations (personal, commercial, informational, political, and spam) was 80 percent on both Twitter and Weibo, with a Krippendorf's alpha of 0.68 on Twitter and 0.72 on Weibo. The percentage agreement of the more precise sub-coding—with two types of personal message, three types of commercial message, three types of informational message, and five types of political message—was 71 percent on Twitter and 68 percent on Weibo, with a Krippendorf's alpha of 0.66 on both platforms. Thus, the percentage agreements and Kripendorff's alphas of both the broad categories and the more detailed sub-codings fall into acceptable levels for this type of research (Lombard et al., 2002). The results of this content analysis are presented in Table 2.

Table 2. Topics of posts by individual users on Twitter and Weibo

	Twitter		**Weibo**		**Significance of difference (Fischer's exact test)**
	Total	**Percentage (%)**	**Total**	**Percentage (%)**	
PERSONAL (of which):	**252**	**50.4**	**106**	**21.2**	**<0.0001**
Sharing information about the self	145	29	96	19.2	0.0004
Personal message to another	108	21.6	10	2	<0.0001
COMMERCIAL (of which):	**121**	**24.2**	**176**	**35.2**	**0.0002**
Entertainment	82	16.4	79	15.8	0.8634
Products and services	34	6.8	97	19.4	<0.0001
Employment	5	1	0	0	0.0619
INFORMATIONAL (of which):	**92**	**18.4**	**169**	**33.8**	**<0.0001**

Inspirational quotations, astrology and life hacks	28	5.6	99	19.8	<0.0001
Viral videos, joke or memes	44	8.8	46	9.2	0.9121
General non-political information	19	3.8	24	4.8	0.5334
POLITICAL (of which):	**34**	**6.8**	**47**	**9.4**	**0.1639**
Expressing an opinion or commenting on society or social practices	16	3.2	15	3.0	1.0000
Sharing news or information or expressing an opinion about current affairs or political events	15	3.0	14	2.8	0.8533
Expressing an opinion on formal political processes	2	0.4	6	1.2	0.4515
Attempting to exert influence on states, companies, organizations or individuals	0	0.0	11	2.2	0.0009
Political humor	1	0.2	3	0.6	0.6242
SPAM	**1**	**0.2**	**2**	**0.4**	**1.0000**
Total	**500**	**100**	**500**	**100**	

The results of this content analysis point to different conclusions about the nature of user speech on these microblogging platforms than previous studies conducted based on trending topics, which have concluded that Twitter is a news medium not a social network (Kwak et al., 2010) and that Weibo is not a news medium but is rather dominated by frivolous content (Asur et al., 2011). At least in individual information production and dissemination, Twitter appears to be used by individuals much more as a social network and less as a news medium. More than half of the posts and retweets by individual US Twitter users either share information about the self or are a personal message to another user. On Weibo, this figure is only 21 percent. Weibo is also very rarely used to send personal messages to another user (two percent of posts), suggesting that users of the platform have a more public orientation. There is no difference in the amount of "frivolous" content (viral videos, jokes, or memes) shared on the platforms or the amount of news, information, or opinions about political or social issues shared on the platforms.

These findings suggest that trending topics are a poor representation of the kind of information that individual users post and forward online. Trending topics are based on hashtags and keywords, include sponsored content, are tailored to user location, and are specifically crafted to show only topics that are new and rapidly growing (Twitter, 2010). The algorithms that control trending topic selection specifically cater to the purposes of the site in presenting a constantly changing array of information that keeps users coming back to check the site frequently; these trending topics should not be seen as a reliable indicator of the topics that individuals actually post about online.

On both platforms, political speech was found to be more frequent than expected based on "the myth of digital democracy"—an idea that was developed before the rise of social media. Political speech makes up almost one in ten posts by individual Chinese Weibo users (9.4 percent). On Twitter, 6.8 percent of posts by individual US-based users were classified as a political speech act. The definition of political speech used, of course, influences the findings. A broad definition of political speech was used here so as not to exclude any speech act that the users would have seen as political or would have been interpreted as political in the context of production or consumption.

However, some research takes a much narrower definition of political speech, requiring it to be active rather than passive, and instrumental rather than informational (see, for instance, Fox, 2013; Gladwell, 2010; Morozov, 2012; Scaff, 1975). The most surprising result we find here is the prevalence of the sub-category of active political speech: i.e. speech which attempts to exert influence on states, companies, organizations, or individuals. Within this random sample of 500 posts by individual users on each platform, 11 posts by Weibo users were "active" forms of political speech that attempted to get other entities to take action to achieve a social or political end. None of the 500 posts on Twitter took this form. This difference is highly significant ($p = 0.0009$, two-tailed Fisher's exact test).

All of the political posts in the Twitter sample were what critics have dismissed as "slacktivist." They shared information or opinions but didn't actually attempt to do anything. In contrast, more than one in fifty posts by individual Weibo users were active attempts to get other users to do something, thereby falling into the category of attempting to exert an influence on states, companies, organizations or individuals. This suggests that although there was not a significant difference between overall levels of political speech on the platforms, Weibo provides important functionalities for active forms of political speech in China in a way that Twitter does not (or does not need to) in the US.

This finding accords with ideas that the Internet may fulfil a more important political function in China because there are fewer offline spaces for political speech and less diverse offline information (Zheng and Wu, 2005), and also with survey data finding that Chinese Internet users report higher levels of online

political activity than US Internet users (Bolsover et al., 2014). The agreement of the results reported in this article with surveys based on user self-reporting suggests that the categories used in this content analysis accord with how individuals understand their own online actions.

It is important to note that this data collection was undertaken after the hardening attitude to freedom of speech online that began in the spring of 2013 after Xi Jinping took office (Benney, 2014; Buckley, 2013), which led prominent political posters to stop posting or to quit Weibo (Moore, 2014). These data cannot speak to the difference between pre- and post-crackdown conditions but they do show that, at least in late 2014, a significant fraction of individual user posts on Weibo concerned political topics—well after some commentators had begun to hail the death of the platform as a space for political speech.

Also of note are the data about the frequency of political humor and expressing an opinion about the formal political process. The small amount of political humor found on both platforms is relatively surprising. A lot of research has pointed to humor as an avenue for political expression enabled by social media, particularly in authoritarian states (Sharbaugh and Nguyen, 2014; Tang and Yang, 2011); however, very few examples of political humor were found in the sample and more traditional types of political speech appear to be much more frequent.

Additionally, these data were collected in the run-up to major political events in both countries—the 2014 US midterm election and the Fourth Plenary Session of the 18th National Congress of the CCP. It was therefore expected that a significant fraction of the political Twitter posts would reference the midterms. However, only two of the 34 political posts on Twitter mentioned the election. We might have expected the 18th National Congress to garner even less attention; important personnel and policy direction announcements are made at these events but there is no element of democratic participation and significant sensitivity surrounds these meeting. Despite this, a greater proportion of political speech on Weibo referred to the National Congress than did Twitter posts about the midterms. While the difference is not statistically significant, it nevertheless suggests that social media may be a much more valuable venue for political speech in authoritarian China compared with the democratic US.

Conclusion

This article presents a new take on the perennial question of the potential of online spaces to act as a venue for political speech and an online public sphere. It has argued that a reason that these questions do not yet have satisfactory answers is that previous research efforts have, when they have used empirical data at all, relied on trending topics or case studies, based on preselected keywords, users, or groups. Thus, questions about the presence of different types of voices in these spaces, and

the prevalence and type of political speech engaged in by users have remained unanswered. The review of the literature also highlighted that the contribution of these platforms to the political process is often seen very differently in authoritarian, as opposed to democratic contexts. Thus, this article has drawn on a random sample of microblog users on Twitter in the US and Weibo in mainland China to address the question of whether these online spaces, which sometimes market themselves as online public spheres, do indeed fulfil these functions.

These data suggest that Weibo appears to provide a space much more akin to a public sphere in China than Twitter does in the US. More than one in ten accounts on Twitter belong to a market-sphere entity, and the colonization of online spaces by commercial forces is one of the main reasons that they do not live up to their participatory potential. In contrast, 98 percent of accounts on Weibo belong to individual users, with almost no market-, civil society- or state-sphere accounts being found in the random sample. This indicates a platform that has not, or at least not yet, been colonized by the market or existing powerholders in society. The idea of Weibo as a space much more akin to an online public sphere is supported by the content analysis of posts made by individual users. Over half of the posts by individuals on Twitter were personal messages, suggesting that the platform functions more as a semi-private social network for individual users rather than as an online public sphere for sharing news and discussing political topics. In contrast, the topic of Weibo posts suggests a much more publically oriented platform.

Levels of political speech were found to be higher than expected on both platforms—9.4 percent of individual users' posts on Weibo and 6.8 percent on Twitter (this difference is not statistically significant). However, a large and significant difference was found in the number of posts that attempted to exert influence on states, companies, organizations, or individuals. On Weibo, one in fifty (2.2 percent of posts) took this form. In contrast, all of the political posts in the random sample on Twitter were of a "slacktivist" nature, sharing information or expressing an opinion rather than encouraging others to take action. This supports previous research, based on survey data, that concluded that Chinese individuals might be drawn to the potential political functions of the Internet because they provide much more valuable affordances as compared to the offline context.

These data also provide a worrying indication of the status of online political speech in the US. Although Twitter markets itself as an online public sphere, a large number of accounts belong to existing powerholders (especially commercial entities), only a very small amount of the political speech in the random sample engaged with the midterm election that fell in the data collection period, and none of the instances of political speech attempted to encourage any form of political action. This supports the critique that online political speech in the US is predominantly "slactivist," and therefore unlikely to influence the balance of power in society.

This study has not, of course, been without its limitations. In using a random sample, it provides quantitative data about user representation and speech without delving into a more precise characterization of who these users are or the nature of these speech acts. This research is also limited in its focus on two platforms during a specific time frame—political speech might be very different on more private platforms, like Facebook or WeChat, or in different time periods, for instance before Xi's crackdown on online rumors or outside periods of formal political events. However, the agreement of these results with previous surveys of Internet users mitigates these limitations.

The limitation of manual geocoding on Twitter, which meant that only 41.6 percent of the 10,007 randomly selected accounts could be geolocated to a particular country, is a limitation that may affect the presented conclusions and their comparability to the Weibo dataset, in which 100 percent of accounts could be geolocated. The hypothesis that this would have affected the comparison of user representation was tested and all of the statistically significant differences reported in this article would hold under this scenario. However, a similar calculation cannot be performed for the coding of the topics posted about by individual users. It might well be the case that individual US-based users who could not be geolocated posted about significantly different topics than those who could be geolocated. Further research efforts could conceivably use network data to infer the location of users who have not entered location data, and thus arrive at a 100 percent geocoded sample to some degree of confidence. However, it would be important to consider whether users of microblog platforms who choose not to enter data that would allow their geolocation have a right not to be geolocated based on data entered by accounts to which they are connected and, thus, whether these network-based efforts to attain higher levels of geocoding on Twitter would infringe the principles of ethical Internet research.

The conclusion that Weibo in China appears to be more akin to a public sphere than Twitter in the US is not intended to suggest that Chinese online spaces are ideal venues for political speech. While a surprising proportion of the posts on Weibo were found to be active forms of political speech, they all fell within permissible topics, such as impoverished elderly, disaster relief, kidnapped children, and abused animals. Additionally, this research considered the ability of these platforms to act as spaces for political speech within the bounds of the existing system. For instance, the content analysis of the spread of voices on the platforms suggests greater domination by existing powerholders on Twitter than on Weibo. However, this does not speak to the overarching system in which the Chinese state exerts much more influence over speech on Weibo than the US state does over speech on Twitter. Taking these wider contexts into consideration, the conditions of Weibo appear much more akin to ideas of an authoritarian public sphere, pushing

gently (and sporadically with more force) at the boundaries of permissibility but largely existing within the constraints of the state.

However, what this comparison does show is that, when the use of microblogs for political speech in the US and China is compared, Weibo appears to be a much better venue for political speech than Twitter. A large number of case studies have shown that Twitter has fulfilled important political functions in the US, however, on a day-to-day basis these data raise questions about its advertised functionality as an online public sphere, and future research should be more critical of the conclusions drawn based on trending topics and case studies. These do not appear to be a good representation of user activity on the platform, the majority of which is personal, or (when it is political) "slacktivist."

Social media do indeed seem to represent a much more valuable contribution to political processes in authoritarian China as opposed to the democratic US, but rather than necessarily conclude that this is because users in democratic states have access to better spaces for political speech offline, further research should critically address the question of whether spaces for political speech in established democracies have been too undermined by commercialization, entertainmentization, and individualization to support the continued functioning of democratic society.

References

Asur, S., Yu, L., Huberman, B., 2011. What Trends in Chinese Social Media. Available SSRN 1888779.

Backstrom, L., Sun, E., Marlow, C., 2010. Find me if you can: improving geographical prediction with social and spatial proximity, in: Proceedings of the 19th International Conference on World Wide Web. ACM, pp. 61–70.

Benney, J., 2014. The Aesthetics of Chinese Microblogging: State and Market Control of Weibo. Asiascape Digit. Asia 1, 169–200.

Bolsover, G., 2017a. Technology and political speech: commercialisation, authoritarianism and the supposed death of the Internet's democratic potential. University of Oxford, Oxford, UK. Available at: https://ora.ox.ac.uk/objects/uuid:f63cffba-a186-4a6c-af9c-dbc9ac6d35fb#permalinkModal.

Bolsover, G., 2017b. Harmonious communitarianism or a rational public sphere: a content analysis of the differences between comments on news stories on Weibo and Facebook. Asian J. Commun. 27, 1–19.

Bolsover, G., 2013. News in China's New Information Environment: Dissemination Patterns, Opinion Leaders and News Commentary on

Weibo (SSRN Scholarly Paper No. ID 2257794). Social Science Research Network, Rochester, NY.
Bolsover, G., Dutton, W.H., Law, G., Dutta, S., 2014. China and the US in the New Internet World: A Comparative Perspective, in: Dutton, W.H., Graham, M. (Eds.), Society and the Internet: How Information and Social Networks Are Changing Our Lives. Oxford University Press, Oxford.
Buckley, C., 2013. Crackdown on Bloggers Is Mounted by China. N. Y. Times.
Burns, A., Eltham, B., 2009. Twitter Free Iran: an Evaluation of Twitter's Role in Public Diplomacy and Information Operations in Iran's 2009 Election Crisis. Presented at the Communications Policy & Research Forum 2009, University of Technology, Sydney, pp. 322–334.
Cammaerts, B., 2008. Critiques on the Participatory Potentials of Web 2.0. Commun. Cult. Crit. 1, 358–377. https://doi.org/10.1111/j.1753-9137.2008.00028.x
Carney, N., 2016. All Lives Matter, but so Does Race: Black Lives Matter and the Evolving Role of Social Media 40, 180–199.
Compton, R., Jurgens, D., Allen, D., 2014. Geotagging one hundred million twitter accounts with total variation minimization, in: Big Data (Big Data), 2014 IEEE International Conference On. IEEE, pp. 393–401.
Davis Jr, C.A., Pappa, G.L., de Oliveira, D.R.R., de L Arcanjo, F., 2011. Inferring the location of twitter messages based on user relationships. Trans. GIS 15, 735–751.
D'heer, E., Vandersmissen, B., Neve, W.D., Verdegem, P., Walle, R.V. de, 2017. What are we missing? An empirical exploration in the structural biases of hashtag-based sampling on Twitter. First Monday 22.
Diamond, L., 2010. Liberation Technology. J. Democr. 21, 69–83. https://doi.org/10.1353/jod.0.0190
Driscoll, K., Walker, S., 2014. Big Data, Big Questions| Working Within a Black Box: Transparency in the Collection and Production of Big Twitter Data. Int. J. Commun. 8, 20.
Express News Service, 2017. Six reasons why Chennai's Jallikattu protest at Marina beach is unique. New Indian Express.
Fox, S., 2013. Is it Time to Update the Definition of Political Participation?Political Participation in Britain: The Decline and Revival of Civic Culture. Parliam. Aff. gss094. https://doi.org/10.1093/pa/gss094
Fu, K., Chau, M., 2014. Use of Microblogs in Grassroots Movements in China: Exploring the Role of Online Networking in Agenda Setting. J. Inf. Technol. Polit. 11, 309–328. https://doi.org/10.1080/19331681.2014.909344

Fu, K., Chau, M., 2013. Reality check for the Chinese microblog space: a random sampling approach. PloS One 8, e58356. https://doi.org/10.1371/journal.pone.0058356
Gerlitz, C., Rieder, B., 2013. Mining one percent of Twitter: Collections, baselines, sampling. MC J. 16.
Ghosh, S., Zafar, M.B., Bhattacharya, P., Sharma, N., Ganguly, N., Gummadi, K., 2013. On sampling the wisdom of crowds: Random vs. expert sampling of the twitter stream, in: Proceedings of the 22nd ACM International Conference on Conference on Information & Knowledge Management. ACM, pp. 1739–1744.
Gladwell, M., 2010. Small change. New Yorker 4, 42–49.
González-Bailón, S., Borge-Holthoefer, J., Rivero, A., Moreno, Y., 2011. The Dynamics of Protest Recruitment through an Online Network. Sci. Rep. 1. https://doi.org/10.1038/srep00197
Gonzalez-Bailon, S., Wang, N., Rivero, A., Borge-Holthoefer, J., 2014. Assessing the Bias in Samples of Large Online Networks. Soc. Netw. 38, 16–27.
Gunaratne, S.A., 2010. De-Westernizing communication/social science research: Opportunities and limitations. Media Cult. Soc. 32, 473–500.
Habermas, J., 1989. The structural transformation of the public sphere: an inquiry into a category of bourgeois society. Polity Press, Cambridge.
Hecht, B., Hong, L., Suh, B., Chi, E.H., 2011. Tweets from Justin Bieber's Heart: The Dynamics of the Location Field in User Profiles, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '11. ACM, New York, NY, USA, pp. 237–246. https://doi.org/10.1145/1978942.1978976
Hindman, M.S., 2009. The myth of digital democracy. Princeton University Press, Princeton.
Howard, P.N., Duffy, A., Freelon, D., Hussain, M., Mari, W., Mazaid, M., 2011. Opening closed regimes: what was the role of social media during the Arab Spring?
Howard, P.N., Hussain, M.M., 2011. The Role of Digital Media. J. Democr. 22, 35–48. https://doi.org/10.1353/jod.2011.0041
Jurgens, D., 2013. That's What Friends Are For: Inferring Location in Online Social Media Platforms Based on Social Relationships. ICWSM 13, 273–282.
Kennedy, S., 2011. Overcoming our Middle Kingdom complex: Finding China's place in comparative politics, in: Kennedy, S. (Ed.), Beyond the Middle Kingdom: Comparative Perspectives on China's Capitalist Transformation. Stanford University Press, Stanford, CA, pp. 3–21.
Kong, L., Liu, Z., Huang, Y., 2014. Spot: Locating social media users based on social network context. Proc. VLDB Endow. 7, 1681–1684.

Krejcie, R.V., Morgan, D.W., 1970. Determining sample size for research activities. Educ. Psychol. Meas. 30, 607–610.
Kwak, H., Lee, C., Park, H., Moon, S., 2010. What is Twitter, a social network or a news media?, in: Proceedings of the 19th International Conference on World Wide Web. pp. 591–600.
Li, R., Wang, S., Chang, K.C.-C., 2012. Multiple location profiling for users and relationships from social network and content. Proc. VLDB Endow. 5, 1603–1614.
Lilleker, D.G., Thierry, V., 2013. The Internet in campaigns and elections, in: Dutton, W.H. (Ed.), The Oxford Handbook of Internet Studies. Oxford University Press, Oxford.
Liu, R.（刘锐）, 2011. The role of opinion leaders in networks of public opinion issues （意见领袖在网络突发公共事件舆论中的作用）, in: Xie, Y.(谢耘耕) (Ed.), New Media and Society (新媒体与社会). Social Sciences Academic Press (社会科学文献出版社), pp. 179 – 211.
Lombard, M., Snyder-Duch, J., Bracken, C.C., 2002. Content analysis in mass communication: Assessment and reporting of intercoder reliability. Hum. Commun. Res. 28, 587–604.
Lovink, G., 2011. Networks without a cause: a critique of social media. Polity, Cambridge, UK ; Malden, Mass.
Moore, M., 2014. China kills off discussion on Weibo after internet crackdown [WWW Document]. URL http://www.telegraph.co.uk/news/worldnews/asia/china/10608245/China-kills-off-discussion-on-Weibo-after-internet-crackdown.html (accessed 12.22.15).
Morozov, E., 2013. To Save Everything, Click Here: Technology, Solutionism, and the Urge to Fix Problems that Don’t Exist. Allen Lane, London.
Morozov, E., 2012. The Net Delusion: The Dark Side of Internet Freedom. PublicAffairs.
Morstatter, F., Pfeffer, J., Liu, H., 2014. When is it Biased? Assessing the Representativeness of Twitter’s Streaming API. ArXiv14017909 Phys.
Nip, J.Y., Fu, K., 2016. Challenging Official Propaganda? Public Opinion Leaders on Sina Weibo. China Q. 225, 122–144.
Penney, J., Dadas, C., 2013. (Re)Tweeting in the service of protest: Digital composition and circulation in the Occupy Wall Street movement. New Media Soc. https://doi.org/10.1177/1461444813479593
Poell, T., Kloet, J.D., Zeng, G., Guohua, 2013. Will the Real Weibo Please Stand Up? Chinese Online Contention and Actor-Network Theory (SSRN

Scholarly Paper No. ID 2304607). Social Science Research Network, Rochester, NY.
Poster, M., 1997. Cyberdemocracy: Internet and the public sphere. Internet Cult. 201, 218.
Rainie, L., Smith, A., Schlozman, K.L., Brady, H., Verba, S., 2012. Social media and political engagement.
Rheingold, H., 1993. The Virtual Community: Homesteading on the Electronic Frontier. MIT Press.
Rout, D., Bontcheva, K., Preoţiuc-Pietro, D., Cohn, T., 2013. Where's@ wally?: a classification approach to geolocating users based on their social ties, in: Proceedings of the 24th ACM Conference on Hypertext and Social Media. ACM, pp. 11–20.
Scaff, L.A., 1975. Two Concepts of Political Participation. West. Polit. Q. 28, 447–462. https://doi.org/10.2307/447366
Shapiro, A.L., 2000. The Control Revolution: How the Internet is Putting Individuals in Charge and Changing the World We Know, 2nd.Printing edition. ed. PublicAffairs, New York.
Sharbaugh, P.E., Nguyen, D., 2014. Make lulz, not war: How online remix and meme culture are empowering civic engagement in the Socialist Republic of Vietnam. Asiascape Digit. Asia 1, 133–168.
Song, F., 2015. State Governance in the Internet Era. Red Flag Manustripts.
Sullivan, J., 2014. China's Weibo: Is faster different? New Media Soc. 16, 24–37. https://doi.org/10.1177/1461444812472966
Sullivan, J., 2012. A tale of two microblogs in China. Media Cult. Soc. 34, 773–783.
Takhteyev, Y., Gruzd, A., Wellman, B., 2012. Geography of Twitter networks. Soc. Netw. 34, 73–81. https://doi.org/10.1016/j.socnet.2011.05.006
Tang, L., Yang, P., 2011. Symbolic power and the internet: The power of a 'horse.' Media Cult. Soc. 33, 675–691.
Twitter, 2010. To Trend or Not to Trend... [WWW Document]. Twitter Blogs. URL https://blog.twitter.com/2010/to-trend-or-not-to-trend (accessed 4.14.16).
Waisbord, S., Mellado, C., 2014. De-westernizing Communication Studies: A Reassessment. Commun. Theory 24, 361–372. https://doi.org/10.1111/comt.12044
Yang, G. (Ed.), 2015. China's Contested Internet. NIAS Press, Copenhagen, Denmark.
Yang, G., 2009. The Power of The Internet in China: Citizen Activism Online. Columbia University Press, New York and Chichester, West Sussex.

Yang, G., Calhoun, C., 2007. Media, Civil Society, and the Rise of a Green Public Sphere in China. China Inf. 21, 211–236. https://doi.org/10.1177/0920203X07079644
Young, L., 2016. Did protests in Zimbabwe really go from ‘tweets to streets’? Wash. Post.
Zheng, Y., Wu, G., 2005. Information Technology, Public Space, and Collective Action in China. Comp. Polit. Stud. 38, 507–536. https://doi.org/10.1177/0010414004273505
Zhou, X., Chan, Y.-Y., Peng, Z.-M., 2008. Deliberativeness of Online Political Discussion. Journal. Stud. 9, 759–770. https://doi.org/10.1080/14616700802207771

Supplementary Online Data Table 1.

Comparison of geolocated Twitter users with all active, non-protected users in the initial random sample

	Geolocated Users (in monthly-active non-protected sample)	**All Monthly-Active, Non-Protected Users**	**Significance of difference**
Number of profiles	4192	10,073	
Average number of profiles followed	506	386	0.12 (unpaired t-test, unequal variances)
Average number of followers	393	324	0.09 (unpaired t-test, unequal variances)
Average number of statuses	3370	2863	<0.01 (unpaired t-test, unequal variances)
Platform Language (only the top five are shown but the whole distribution is used for the significance test)	English (US) – 55.92% Spanish (Spain) – 15.94% Japanese – 6.44% Thai – 4.63% Portuguese – 3.63%	English (US) – 50.00% Spanish (Spain)– 14.42% Japanese – 11.04% Arabic – 4.09% Portuguese – 3.69%	<0.01 (chi-squared test)
Time Zones (only the top five are shown but the whole distribution is used for the significance test)	(None – 41.44%) Eastern Time (US & Canada) – 5.82% Central Time (US & Canada) – 4.94% Pacific Time (US & Canada) – 3.67% London – 2.93% Bangkok – 2.86%	(None – 56.49%) Eastern Time (US & Canada) – 4.57% Central Time (US & Canada) – 3.50% Pacific Time (US & Canada) – 2.43% Tokyo – 2.18% Brasilia – 2.02%	<0.01 (chi-squared test)

Supplementary Online Data Table 2.
Number of users in the geolocated Twitter sample by country

Country	Number of users (per listed country)	Percentage of Twitter users (per listed country, %)	Country	Number of users (per listed country)	Percentage of Twitter users (per listed country, %)
USA	962	22.95	Indonesia	410	9.78
UK	367	8.75	Japan	298	7.11
Brazil	176	4.20	Spain	166	3.96
Turkey	153	3.65	Mexico	137	3.27
Canada	101	2.41	Russia	101	2.41
Argentina	90	2.15	Saudi Arabia	88	2.10
Venezuela	78	1.86	Colombia	74	1.77
France	70	1.67	India	54	1.29
The Netherlands	53	1.26	The Philippines	52	1.24
South Africa	42	1.00	Australia	39	0.93
Chile	38	0.91	Nigeria	36	0.86
Italy and Malaysia	33	0.79	Germany	28	0.67
Egypt	27	0.64	Ireland	26	0.62
South Korea	24	0.57	Peru	22	0.52
Thailand	18	0.43	Ecuador, Kuwait and UAE	17	0.41
Dominican Republic	14	0.33	Guatemala	13	0.31
Pakistan and the Ukraine	11	0.26	Greece, Norway, Poland and Uruguay	10	0.24
New Zealand, Panama, Paraguay and Puerto Rico	9	0.21	Czech Republic, El Salvador and Portugal	8	0.19
Belgium, China, Denmark, Serbia and Taiwan	7	0.17	Belarus, Kenya, Singapore and Switzerland	6	0.14

Honduras, Iran, Morocco and Sweden	5	0.12	Bangladesh, Finland, Ghana, Lebanon, Palestine and Qatar	4	0.10
Azerbaijan, Croatia, Cyprus, Iraq, Israel, Kazakhstan, Libya, Nepal, Romania and Vietnam	3	0.07	Algeria, Austria, Bahrain, Bolivia, Bulgaria, Hong Kong, Jamaica, Jordan, Latvia, Moldova, Montenegro, Suriname, Tanzania, Tenerife, Tunisia and Uganda	2	0.05
Afghanistan, Albania, Angola, Bahamas, Belize, Benin, Cameroon, Cayman Islands, Cote d'Ivoire, Cuba, Estonia, Fiji, Gabon, Hungary, Kurdistan, Liberia, Macedonia, Mozambique, Oman, Papua New Guinea, Rwanda, Senegal, Slovenia, Somalia, Sri Lanka, Sudan, Swaziland, Trinidad and Tobago, Yemen, Zambia and Zimbabwe	1	0.02			

Supplementary Online Data Table 3.

Summary of Post Characteristics by Individuals on Twitter and Weibo over a Four-Week Period

	Twitter	**Weibo**	**Significance of difference**
Number of users who posted at least once in the four-week time period	546	550	
Number of posts	43,543	8907	<0.01 (unpaired t-test, unequal variances)
Average number of posts per individual per week	20	4	<0.01 (unpaired t-test, unequal variances)
Number of retweets/forwards	17,076	5107	<0.01 (unpaired t-test, unequal variances)
Percentage of posts that are retweets/forwards	39.2%	57.3%	<0.01 (Chi-squared test with Yates correction)
Number of retweets/forwards from verified accounts	3528	2568	<0.01 (Chi-squared test with Yates correction)
Percentage of individual users' posts that are retweets/forwards from verified accounts	8.1%	28.8%	<0.01 (Chi-squared test with Yates correction)
Percentage of retweets/forwards that were originally posted by verified users	20.7%	50.3%	<0.01 (Chi-squared test with Yates correction)

Supplementary Online Data Table 4.

Example Posts Representative of Topical Coding Scheme Categories

The example posts presented below are based on example posts from the dataset that were coded in each of these categories. However, text of these posts has been altered, while maintaining their original meaning, to maintain the anonymity of the users who were included in the data collection. The decision to present Weibo posts only in their English translation is part of this process of protecting the anonymity of these users. Due to the fact that Chinese does not use spaces between words. Due to this fact, spaces are added to Weibo usernames when they are translated to English in the below table and the usernames enclosed in quotation marks.

	Twitter	**Weibo**
PERSONAL (of which):		
Sharing information about the self	**Original:** *I shouted at my food in the oven*	**Original:** I was singing a song, singing, singing then I started crying, why?!?
Personal message to another	**Original:** @jaydenjames tell your mom i ain't shy i just don't like other guys around my woman lol dont tell her	**Original:** Breaking up is always horrible "@Li Junwen"
COMMERCIAL (of which):		
Entertainment	**Original:** A minute into the episode and I think I'm going to be sick #TheWalkingDead	**Forwarded:** [The eight most important theme parks in China opening in 2015] The eight most important theme parks opening across China in 2015, include Shanghai Disneyland, Hello Kitty Holiday Park, Lotte World, Eastern Hollywood Tyrants and Stars' Westward… you can visit world-class theme parks without going abroad, which one are you mist excited about? **Forwarding user added comment: "**Witch Sarah Bobo" "@xwppp_1842" "@Eat, then sleep" "@Thousand-year eggs and fried potato strips child" What about it? [emoji for stealthy music] [emoji for stealthy music]
Products and services	**Original:** Only bought Addidas shoes tomrw were getting more clothing ;)	**Original:** The seafood cakes were incredibly fresh [emoji for gluttony]. I love the cook who made them [emoji for love you]
Employment	**Original:** Sara is visiting the office Friday to show us some new products they're going to release next year [four emojis of confetti streaming out of a party hat]	**Original:** Do you want to make money like a Korean snack importer? [emoji for money] Join my group! [twinkling stars emoji] Don't ask silly questions. I like to do business with smart people. Do you want to make a lot of money selling skincare marks? Send me a private message and we can chat. [emoji for smiling face] [User supplies their Weixin/Wechat ID number]
INFORMATIONAL (of which):		
Inspirational quotations,	**Original:** It's not about how nice a person you are, it's about how nice a person you try to be.	**Forward:** There will always be someone who can easily achieve what you have worked very hard for for a long time.

astrology and life hacks		
Viral videos, joke or memes	**Retweet:** RT @SchoolDays: why does my teacher always draw donuts on my work? [link to image of a red zero within a red circle drawn on lined paper]	**Forward:** I went to the hospital to protect my girlfriend… Laughing and crying haha haha [attached image of a long, humorous story who meets a mugger while walking at night with his girlfriend] (The conversion of long text to images that are then uploaded as attachments is very common on Weibo.) **Forwarding user added comment:** Am I also this kind of person? "@Fallen oak leaves"
General non-political information	**Retweet:** RT @t3ftoit: How Stress Makes Us Lose Sight of Our Goals [link to offside article on site Live Science]	**Forward:** [The Nine Principles of Effective Powerpoint Presentations] College student job hunting season has his the peak! Are you starting to look for a job? Do you feel that you are already to smart to learn? Quickly learn the necessary powerpoint presentation skills for the workplace! 1) 10-20-30 principle: there powerpoint should not exceed 10 slides, the speech should not exceed 20 minutes and the typeface should not exceed 30. 2) Make it interesting 3) Slow down 4) Make eye contact 5) 20-20 principle: 20 slides, 20 seconds of speech each. Must forward!
POLITICAL (of which):		
Expressing an opinion or commenting on society or social practices	**Retweet:** RT @LaurentSim: Pro-life, where a white embryo is more important than the lives of black children and young people	**Original:** Once you start working, you realize how hard it is to make money. There is a reduction in compassion. Already there is very little given to beggars and compassion for those busking on public transport. I don't know if this is a good thing or a bad thing? [emoji for thinking]
Sharing news or information or expressing an opinion about current affairs or political events	**Retweet:** RT @usacsmret: 50 million on food stamps. 12 million on unemployment. 5 million on welfare. Obama's "recovery" is worse than Bush's reces…	**Forward:** [Do you know how many domestic banks there are in China! (the most complete list of financial institutions)] 68 trusts, 91 investment funds, 67 foundations, 111 securities companies, 832 banks, 115 finance services companies, 40 financial lenders, 265 financial management companies. **Forwarding user added comment:** This data is too old; 184 financial services companies have been approved. "@Professor Lu Mintai" The statistics on financial lenders must be incorrect; who can help? "@Zhang Shaoxin" retweeted status.
Expressing an opinion on formal political processes	**Retweet:** RT @HelloJamesDean This whole election process would be much better if we changed to a "Hunger Games" format	**Forward (made by a professor of constitutional law):** I believe that after the 18th CCP Fourth Plenary Session, the process of constructing the Constitutional Supervision System will definitely experience substantial progress because now the country's highest leadership has emphasized the importance of forcefully tackling difficult problems (踏石留印抓铁有痕- "tread stone, leave mark; grab iron, leave scar") to achieve practical results. Achieving practical progress in this vein must begin by tackling two problems: formulating a constitution to control procedural law and setting up constitutional supervision of specialized agencies. Even if a temporary constitutional court is not

		established, at least a guardian council could be set up inside the National People's Congress. **Forwarding user added comment:** A theoretical breakthrough, we need great wisdom!
Attempting to exert influence on states, companies, organizations or individuals	[No posts of this type were present in the Twitter dataset.]	**Forward:** #Jinggu 6.6 magnitude earthquake# Banpo Village People's Government: on the night of the earthquake, there were just a few sloppy groups looking at state housing. The houses belonging to regular people were ignored. The third day after the earthquake, still no relevant personnel have paid a visit. We are 40 kilometres from the epicentre of the earthquake. "@ Yunnan Radio and TV" "@Urban barcode" "@Yunnan People's Livelihood Channel" "@Phoenix Weibo News" "@Spring City Evening News" "@Phoenix Television" (The user tags six regional media outlets and attaches a photo gallery of eight photos shows the damage that the earthquake has done to their house). **Forwarding user added comment:** #Jinggu Earthquake# No one is paying attention to the people on the streets. No one has come to help. There are no tents. My neighbour is pregnant and the whole family is sleeping outside. Temperatures are lower in the countryside. The only people who have tents are those who work at the school. Other people are sleeping outside. The government must help the people (为人民服务 - a CCP political slogan) "@Yunnan People's Livelihood Channel" "@Urban barcode"
Political humor	**Retweet**: RT @TedOfficialPage: Freshmen year vs senior year. (attached photo *shows then U.S. president Obama's looking young and excited at the start of his term and older and worried at the end of his term)*	**Forward:** After looking at North Korea's fat boy leader… [emoji for tears rolling down face] (link to a humorous story of a round-based battle between major countries) (the phrase used in this tweet 三胖, literally three fat, is commonly used by Chinese netizens to refer to Kim Jong-un the third in the line of heredity leaders of North Korea who are seen as notable for being overweight in a country in which the majority of people suffer from hunger and malnutrition) **Forwarding user added comment:** It's the era of the ambitious and ruthless Korean Kim Jong-un (the name 三胖, the third fat one, is again used to refer to the North Korean leader)"@writer Zhang Weihuo" [laughing emoji] [laughing emoji] [laughing emoji] [emoji of the English word good]
SPAM	**Original:** Retweets and quotes count. If you see this quote and retweet, I'll retweet your quote. #BieberBiggestFans #vote4us #battleofthebands	**Forward:** Forward and follow this microblog. After 15 hours, the platform will select and send on user ￥500 yuan. (approximately £50)