Using New Communication Technologies to Study Contemporary China

by

Charles Chang

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Environment and Resources)

at the

University of Wisconsin-Madison

2016

Date of final oral examination: 8/9/2016

The dissertation is approved by the following members of the Final Oral Committee:

Melanie Manion, Professor, Political Science
Scott Gehlbach, Professor, Political Science
Pierre Landry, Professor, Political Science
Qunying Huang, Assistant Professor, Geography
A-Xing Zhu, Professor, Geography
Harvey Jacobs, Professor, Gaylord Nelson Institute

Table of Contents

I.	INTRODUCTION	1
II.	MAPPING COMMUNITIES IN RAPIDLY DEVELOPING COUNTRIES BY USING MEDIUM-RESOLUTION REMOTELY SENSED IMAGERY AND GEOTAGGED INFORMATION	16
	1. Introduction	16
	2. Study Area and Datasets	18
	3. Method	21
	4. Results in Visual Representation	29
	5. Assessments	33
III.	PUTTING CENSORSHIP IN ITS PLACE: GAUGING BOUNDARIES OF AUTHORITARIAN CONFIDENCE BY ANALYZING GEOTAGGED POSTS IN CHINESE SOCIAL MEDIA	49
	1. Introduction	49
	2. Sina Weibo	52
	3. Data	55
	4. Identifying Posts for Analysis	57
	5. Analysis	59
	6. Conclusion	
IV.	NETIZEN CONFIDENCE IN AUTOCRAT DISPENSED INFORMATION OVER AN INTRUSIVE EVENT	92
	1. Introduction	92
	2. Theoretical Framework	94
	3. Hypothesis	96
	4. Empirical Strategy	100
	5. Findings	111
	6. Conclusion	117

List of Tables

II. MAPPING COMMUNITIES IN RAPIDLY DEVELOPING COUNTRIES BY USING MEDIUM-RESOLUTION REMOTELY SENSED IMAGERY AND GEOTAGGED INFORMATION

Table 1. The Landsat TM/ETM Images Used in this Study 39

Table 2. Data Description for GIS Data other than Remotely Sensed Imagery 40

Table 3. Major Types of Land Covers and Urban Land Use Categories 41

Table 4. Confusion Matrix of Urban Land Use Major Categories Accuracy Assessment 42

III. PUTTING CENSORSHIP IN ITS PLACE: GAUGING BOUNDARIES OF AUTHORITARIAN CONFIDENCE BY ANALYZING GEOTAGGED POSTS IN CHINESE SOCIAL MEDIA

Table 1. Four Political Catalysts, 2014−2015 79

Table 2. Critical Posts and Distance to Tiananmen Square, Logistical Regression Models 83

IV. NETIZEN CONFIDENCE IN AUTOCRAT DISPENSED INFORMATION OVER AN INTRUSIVE EVENT

Table 1. Number of comments on Sina News and posts on Sina Weibo both before and after the Attack 120

Table 2. Political, Non-Political Netizens and the Transformation between them 5 Months before and 5-days amid the Kunming Railway Station Attack 121

Table 3. Netizen Movements around Kunming Railway Station before and after the Attack 122

Table 4. Netizen Movements around the Dashuying Uighur Community before and after the Attack 123

Table 5. Netizen Movements around Kunming Railway Station Differentiated by Connections to Government or Gated Community 124

Table 6. Netizen Movements around the Dashuying Uighur Community Differentiated by Connections to Government or Gated Community 125

List of Figures

II. MAPPING COMMUNITIES IN RAPIDLY DEVELOPING COUNTRIES BY USING MEDIUM-RESOLUTION REMOTELY SENSED IMAGERY AND GEOTAGGED INFORMATION

Figure 1. Study Area 43

Figure 2. Workflow of Land Use Classification 44

Figure 3. Layout and Views of Typical Land Uses 45

Figure 4. Land Cover Change in three Sample Areas 46

Figure 5. Maps of Urban Land Use in Major Categories and Sub-Categories 47

Figure 6. Chord Diagram of Classification Accuracy in Residential Sub-Categories 48

III. PUTTING CENSORSHIP IN ITS PLACE: GAUGING BOUNDARIES OF AUTHORITARIAN CONFIDENCE BY ANALYZING GEOTAGGED POSTS IN CHINESE SOCIAL MEDIA

Figure 1. Frequency of Geotagged Political Posts, Beijing, July 1, 2014−June 15, 2015 77

Figure 2. Spatial Distribution of Political Posts, Beijing, July 1, 2014−June 15, 2015 78

Figure 3. Spatial Distribution of Political Posts, Beijing, Four Political Catalysts 80

Figure 4. Spatial Distribution of Critical Political Posts, Beijing, Four Political Catalysts 81

Figure 5. Spatial Distribution of Critical Political Posts after Release of Under the Dome, Beijing 82

Figure 6. Critical Posts and Distance to Tiananmen Square 84

IV. NETIZEN CONFIDENCE IN AUTOCRAT DISPENSED INFORMATION OVER AN INTRUSIVE EVENT

Figure 1. The theoretical Framework of Information Contestation between Autocrats and Netizens 126

Figure 2. Number of Political Posts Sent by Hour over the Event 127

Figure 3. The Avoidance of Network Distance Illustrated near Kunming Railway Station and the Dashuying Uighur Community 128

List of Appendices

III. PUTTING CENSORSHIP IN ITS PLACE: GAUGING BOUNDARIES OF AUTHORITARIAN CONFIDENCE BY ANALYZING GEOTAGGED POSTS IN CHINESE SOCIAL MEDIA

Appendix A. Data Collection 85

Appendix B. Differences between Geotagged and Non-Geotagged Posts 88

IV. NETIZEN CONFIDENCE IN AUTOCRAT DISPENSED INFORMATION OVER AN INTRUSIVE EVENT

Appendix A. Data Collection 133

Appendix B. Differences between Geotagged and Non-Geotagged Posts 136

Appendix C. Information manipulation on media and social media, including Sina News and Weibo 141

Appendix D. The Space of government and gated community 144

Acknowledgements

In completing this work, I should like to express my thanks to Melanie Manion for her kind initial prompting to undertake this dissertation topic, her warm encouragement and exacting conversations, and her diligent mentorship. I feel most fortunate that I met her during my inconclusive peregrinations. Also, I am most grateful to all my committee members—their generous help and consideration make me feel all the hard work worth the effort.

I also thank my adoptive father, Yi-Fu Tuan, for his unconditional love and support. And I am forever indebted to my friends at UW-Madison and China who made my graduate student life happy, fulfilling, and exciting!

Abstract

New communication technologies (internet, social media, smartphones) have flooded societies in most parts of the world, altering social and political relations in them. This transformation is of great importance to political scientists who want to know, for example, how in a non-liberal democratic regime the power relationship between rulers and ruled shifts once the ruled possess the new means of communication. Will netizens with technical communicative savvy succeed in destabilizing the regime or will the regime gain even firmer control by using the same technologies to greater effect? To get answers, I examine the contestation between the most technically communicative-savvy netizens—smartphone social media users—and the most technically competent autocrats—Chinese Communist Party rulers. My dissertation consists of three parts, held together by a novel "view-from-above" approach that complements the boots-on-the-ground approach traditionally used by social scientists. Part One introduces such a method—one that involves the use of smartphone social media—to construct a detailed land use map when such maps are not available in a developing country or are available but are not accessible in a non-democratic one. Part Two makes full use of smartphone social media to gauge boundaries of regime confidence: how Chinese rulers permit unorganized clustering of political talk at sensitive times and places. Part Three makes full use of smartphone social media to gauge the degree of netizen confidence in the information rulers dispensed over an intrusive event: how netizens are engaged with rulers' information but avoid places due to distrust. In the analysis in all three parts, I find myself drawn repeatedly to the role of place—that it resonates, that it is vocal, and that it can even be a megaphone of the political message. Place and politics is also a likely future direction for my research.

Introduction

New communication technologies (internet, social media, smartphones) have flooded societies in most parts of the world, altering social and political relations in them. This transformation is of great importance to political scientists who want to know, for example, how in an authoritarian regime the power relationship between rulers and ruled shifts once the ruled possess the new means of communication. Will netizens with technical communicative savvy succeed in destabilizing the regime or will the regime gain even firmer control by using the same technologies to greater effect? Do netizens trust autocrats when they take steps to counter an exogenous shock such as a natural disaster or a terrorist attack? What political views do netizens in urban communities hold? What is the likelihood that these views lead to collective action? Is messaging across distance sufficient for that purpose or is face-to-face contact a necessity? If the latter, we will need to know the size of the community and its distance from others of similar affluence and degree of political awareness. To get answers, a map is required. A map is also required when we want to have a sense of the extent to which political forces are at play as urban buildings expand into farmland and as one kind of community expands into another kind. A study of such expansions and transgressions in just one part of the city cannot give us an idea of the full political force exercised in the entire city, much less in a constellation of them.

I have raised a number of issues here, which political scientists typically address by interview and survey methods. My methods complement the boots-on-the-ground approach with a map that is essentially “a view from above.” Looking at such a map is like seeing the ground from the air, with sufficient detail to almost enable us to see socio-political relations over space at a glance. My methods also have this spatial, “seen from above” characteristic in that they

depend on big data in cyberspace—"big" as compared with the data generated by interviews and surveys. Whereas "map" suggests the concrete, depicting a level of reality almost like what we see from an airplane, big data suggest a high degree of abstraction, so much so that the reality they supposedly enlighten almost fades from view: the pyrotechnics of analysis become almost an end in itself. The big data I use, by contrast, hew much closer to the real world for they are the sayings (the "puts," in technical language) of real geotagged human beings.

My dissertation consists of three parts, held together by the common methodology introduced above. Part One introduces a technique—one that involves the use of social media—to construct a detailed land use map when such maps are not available in a developing country or are available but are not accessible in an authoritarian one. Part Two makes full use of social media to gauge boundaries of authoritarian confidence. Part Three makes full use of social media to gauge the degree of netizen confidence in the authoritarian government. In the analysis in all three parts, I find myself drawn repeatedly to the role of place—that it resonates, that it is vocal, and that it can even be a megaphone of the political message. Place and politics is also a likely future direction for my research.

Part One: The View from Above

In Part One, my task is to discover what communication patterns, especially those of a socio-political nature, are in urban space. These patterns are probably known to the Chinese Communist Party, given its tenacious surveillance of citizens. But they are unknown to ordinary people and are largely inaccessible to researchers without special clearance. I propose a method that will overcome this lacuna and use Kunming, a middle-size Chinese city, to illustrate the procedure. The procedure begins by taking up change at both the periphery and the core of the city. Mapping peripheral expansion from a non-built-up surface (e.g., agriculture) to a built-up

surface is relatively simple because the expansion so mapped seldom involves urban renewal. Put another way, it is largely horizontal. Where the expansion is largely horizontal, the technique of remote sensing is adequate to the purpose. Mapping expansion at the urban core is, however, far more challenging, for much of it is from small low-lying, old buildings to big, new tall ones; it is, in a word, vertical. In this case, remote sensing technique is inadequate and a new method is needed. The new method depends on big data, which has the power to disentangle urban complexities, one of which is to identify specific land use. Whereas the much-used tool of remote sensing can discern built-up areas from forest, grassland, and water, it cannot show whether the built-up area is an airport, a commercial district, or a residential compound, distinctions that are crucial to understanding the urban environment. Needed, then, are other sources of information. These include, besides remotely sensed images taken at dense time intervals, place-specific data, known as points of interest (POIs) that online mapping services provide; locational data embedded in conversations conducted through smartphones and other social media; and information from such sources as phone prices, road networks, and online real estate listings. For annual classification of built-up areas, accuracy achieves 78% in the urban periphery; cumulatively, accuracy reaches 90% by 2013, the end of my research period. As to the urban core, I took further steps, including the use of information concerning the age of buildings in online real estate listings to determine the approximate year when renewal occurred.

Having mapped both the periphery and the core in urban expansion, my next step is to identify the major land use categories in the built-up areas. Following common usage, these are industrial, commercial, office, transportation, recreational, and residential. The procedure I adopt is to consider land parcels: small pieces of land developed from non-built-up to built-up during the same year and which may therefore have the same land use. However, the procedure cannot

discriminate between adjacent land parcels on two sides of a road that have different land uses and are yet built in the same year; nor can it discriminate across land uses in the urban core in which the development is from built-up areas to differently built-up areas within my study period. To overcome this difficulty, I make use of the current road network and the information provided by nearby POIs, which I aggregate to provide a hierarchy based on size, namely, building complex, individual building, and sub-building. To give some examples, an airport falls under the major category of transportation and yet the building complex may well have smaller uses, such as cafés and newsstands. Similarly, a museum falls under the recreational category but may have a store selling souvenirs. A sub-building is single-use, such as shoe store. As to major categories that are underrepresented by POIs, one way to address this problem is to monitor the temporal pattern of geotagged social media posts over each land parcel, because the pattern sent from homes is distinctive. Within a 24-hour period on weekdays, more posts are likely to be sent very early in the morning and late at night. Comparing the posts sent in the weekdays to the weekends, more posts are sent from 9 to 5 during weekends. Generally, more posts are sent during holidays and special family-centered occasions than during non-special days.

Using my method, I am able to obtain results in land use classification of the major categories with acceptable accuracy. Overall, it reaches 74%. Variation of accuracy is, however, large from one category to the other. Thus, in commercial, office, transportation, storage and farm houses, and residential categories, accuracy reaches levels of 88%, 67%, 66%, 69%, and 84% respectively. This is satisfactory, but in recreation and manufacturing, it is lower. I present a map that shows the major categories of land use in both the downtown and suburban areas of Kunming. It reveals a strikingly high degree of fragmentation.

I now choose for consideration one major land use category—residential. I subdivide the

residential category into gated communities, ordinary communities, and urban slums, based on socioeconomic status of residents. Socioeconomic status is of interest to me because it has political implications. But, before I can explore the political implications, I need to meet the challenge of delineating the residential sub-categories accurately, for they often contain nonresidential uses. One way is to use POIs and online real-estate listings. This latter gives information about basic type of construction, whether the buildings are villas or high-rises; I use this information to separate different types of communities. This aggregate information can then be supplemented by POIs, which indicate the type and number of surrounding services: for example, gated communities have fewer inclusive services than ordinary communities. The pattern of social media posting differentiates the residential sub-categories and, in particular, urban slums from the other two because urban slums lack POIs and real estate listings. Lastly, the price of mobile devices is a gauge of resident wealth and so indicates whether the man or woman who uses them lives in a gated community, an ordinary community, or an urban slum.

As to mapping accuracy, my method yields the following results: gated communities at 73%, ordinary communities at 72%, and urban slums at 89%. These results and the distribution of the three types of community are mapped and shown in Part One. Remotely sensed imagery, as noted earlier, cannot tell change within the urban core, for change there is in a vertical direction. It also cannot clearly distinguish gated communities from ordinary communities in the urban core. My integrated method can do both. In fact, the urban core offers a special advantage in mapping accuracy for it has a large number of POIs that are highly indicative of land use. Moreover, its social media are both exceptionally abundant and varied. These indicators, when combined, enable me to separate gated communities from ordinary communities where the boundaries are unclear. Unlike these two types of communities, urban slums are shapeless and

fragmented. They do not fit into zones. Poor or otherwise marginalized people take over whatever space has not yet been occupied or is considered undesirable.

My method ends in a map. What the map does for the researcher is that it uniquely empowers him or her to see, at a glance, the changing size, location, and configuration of urban political phenomena. Understanding such phenomena can be relatively straightforward—for example, the infringement of an expanding urban area on agricultural land. But the visual form of evidence (i.e., maps) can also raise complex and fruitful questions of a sociopolitical nature. Consider the striking fragmentation of land use I show in Kunming in major categories and sub-categories. That a centralized and authoritarian government should have allowed the city of Kunming to expand in a seemingly haphazard manner is surprising if only because it is so much at odds with the rigidly planned Chinese cities of the past. Or is this simply a striking case of economic imperative trumping political ideology? I show, for example, that the three types of communities—gated, ordinary, and slum—are all fragmented. In particular, urban slums are a veritable patchwork of shabby housing distributed throughout Kunming, sometimes smack against ordinary communities or even fancy gated ones. Could the government have allowed (and perhaps even promoted) fragmentation in order to make mass collective action difficult? For collective action depends on effective communication, which has to be more than just exchanges of social media; it also is facilitated by body language and bodily contact of the sort that occurs in daily life—and it is bonding at this bodily level, which can only happen when people live in the same area, that is disrupted by the city's excessive parceling and fragmentation. The maps that my method produces are not the usual tools of political science, and yet I hope to have shown that they stimulate new questions of keen interest to the discipline. To reiterate: the map's unique power to raise new questions lies in its presentation of broad

spatial relationships at a glance, an image of the real world. In this regard, statistical tables and graphs, the modes of representation commonly found in the scholarly papers of political science, seem deficient.

Part Two: Social Media in Place—Authoritarian Confidence

The second part of my dissertation, a joint paper with my advisor, gauges boundaries of authoritarianism. A mere decade ago, the technological revolutions that made real-time exchange of information and opinions among ordinary citizens possible seemed to pose a serious challenge to authoritarianism. Yet, even as "pundits, policymakers, and the popular press" noisily proclaimed a social media triumph over clumsy dictators, the authoritarian response was evolving from blunt suppression to a smarter mix of censorship, surveillance, and propaganda. Political science, while never a main force of cyber-utopianism, now better grasps how authoritarian rulers can and do exploit new information and communication technologies to promote a fundamental interest in political stability. Moreover, our scholarly understanding of how "networked authoritarianism" works in different authoritarian contexts has grown, to take account of the new information management.

Nowhere is the authoritarian machinery to discover, shape, and censor online public talk more developed than in China, under a highly adaptive communist single-party regime. The world's most advanced national firewall selectively blocks access to foreign content. A cyberpolice of perhaps 30,000 carries out fulltime surveillance. Hundreds of thousands of paid propagandists flood the cyberspace with political messages that redirect rumors and amplify regime support. These human armies complement the operation of dozens of software products that filter and delete "harmful" information, viewpoints, images, and language. With this machinery, China's rulers manage a huge online population: 668 million netizens in mid-2015,

with 89 percent of them accessing the internet on smartphones.

We analyze temporal and spatial clustering of some 6.7 million geotagged posts on Sina Weibo, China's sophisticated version of Twitter, to answer the following question: How do powerful authoritarian rulers, obsessed with stability, manage netizens who engage in online political talk (including critical talk) and happen to cluster in a common physical space? In particular, how sensitive are rulers to potentially mobilizing focal points?

The Chinese state is experienced and capable at managing actual outbursts of unsanctioned grassroots collective action—although it evidently prefers to prevent them. Our point of departure is the novel research of King, Pan, and Roberts that finds Chinese censorship of the social media aims not to prevent criticism but to prevent collective action by deleting netizen talk about its actual occurrence, past or future. Put differently: Censorship of Chinese cyberspace reacts to netizen attention to where other netizens (and non-netizens) organize themselves politically to locate in a common physical space. This finding is consistent with a view in the literature that autocrats most fear a coalition of "like-minded people" gathering together in some place to challenge "the regime's stability mantra." King, Pan, and Roberts arrive at their conclusion through analysis of social media content, but we see its clear implications for the role of *place* in social media censorship. Conceptually and methodologically, we walk through the door they open. We leverage visual and statistical discernibility of spatial clustering of posts (including critical posts) about politics, posted in Beijing in 2014−2015, to measure Chinese official sensitivity to unorganized clustering of like-minded ordinary netizens, neither activists nor celebrities, in physical space. We use this measure to go a step further in gauging boundaries of authoritarian confidence about political control of Chinese society.

Analytically, we treat visual and statistical discernibility of spatial clustering of political

posts as an indication that the like-minded netizens of interest to us do cluster spatially, but this point is not in itself of particular analytic interest. Rather, if we can observe that political posts in our dataset of already censored posts discernibly cluster, it interests us because of implications about (absence of) censorship of netizens who coincide in physical space and gather in cyberspace to talk about politics. In particular, we look in our analysis for spatial clustering of political posts, including critical political posts on politically sensitive dates.

Given the official obsession with social stability, the sensible theoretical place to start is a claim of observed regime *vulnerability.* That is, China's rulers sense a regime threat in the unorganized clustering in physical space of netizens talking about politics; blunt censorship of Sina Weibo political posts reflects this obsession. We formulate our null hypothesis accordingly, to reflect censorship. Indiscernibility of spatially clustered political posts in our dataset of already censored posts leads us to reject the null in favor of an alternative hypothesis, a *confident* regime.

We analyze different situations and catalytic political events to test progressively stricter versions of the null. Variation in discernible spatial clustering of political posts across situations and political catalysts is our way of mapping observed vulnerability of Chinese rulers vis-à-vis the society they govern. It allows us to make inferences about regime sensitivity to focal points that (in principle) have greater than usual potential to mobilize public opinion, perhaps even collective action. We are especially interested in events that seem neither obviously threatening nor obviously unthreatening. For example, we analyze sensitivity to netizen spatial clustering after release of Chai Jing's *Under the Dome*. Places themselves can also be focal points. For example, we analyze sensitivity to netizen clustering at Tiananmen Square.

Even after several catalytic events, we find visually and statistically discernible clustering

of netizen political posts, including critical posts, but no criticism posted at Tiananmen Square. We conclude that Chinese rulers are confident enough to permit unorganized spatial clustering of political talk on Sina Weibo, the most potentially subversive form of Chinese social media. This confidence extends to most politically sensitive times and politically sensitive places. However, some places (Tiananmen Square, for example) are focal points that resonate politically with rulers and netizens alike. Especially at politically sensitive times, these places themselves function as (censurable) media.

Research by King, Pan, and Roberts shows that Chinese censorship in cyberspace aims to delete netizen talk about collective action. In this part of our research, we bring geography into the study of social media and show how physical space matters for information management in authoritarian regimes. We find that China's rulers permit clustering of critical talk by ordinary netizens, even if the talk is sometimes subversive, provided it does not occur at sensitive times or sensitive places. On the extremely sensitive anniversary of June 4, for example, we find Beijing netizens are either not permitted or do not dare to voice discontent. In other examples, such as the investigation into the corruption of mega-tiger Zhou Yongkang, the release of Chai Jing's antipollution documentary *Under the Dome*, and the regime-sponsored Nanjing Massacre Memorial Day, we find netizens in Beijing do voice discontent—but, as our analysis shows, their discontent clusters at some distance from Tiananmen Square. What is remarkable about Tiananmen Square at politically sensitive times is that even without netizen voices of discontent or the presence of a single netizen, it still has the power to be a medium, indeed a megaphone, that China's authoritarian rulers evidently find threatening.

Part Three: Social Media in Place—Netizen Confidence

On Saturday, March 1, a group of knife-wielding Uighurs attacked the Kunming railway

station. The state media quickly proclaimed this attack an act of "terrorism," and almost as quickly—only three hours later—it claimed the event "under control." Following these announcements, the state media, along with its commercial partners, rolled out a series of stories that dominated the entire media landscape for a few days. The stories were sensational: some were about heroic policemen risking their life battling the terrorists; others told of the severely wounded who lived a nightmare they couldn't shake off. No doubt there were images of state officials warmly shaking the hands of the victims. As for civil society—celebrities among Uighurs, Tibetans, and athletes acknowledged by the state—all condemned the terrorists and vowed to stand by law and order. It would seem that Chinese autocrats, unlike politicians in a democratic country, were decisive and efficient in confronting terrorism and, by extension, any intrusive event. But is this anything like the true course of events?

Governments, democratic and authoritarian, have to deal with intrusive events. However, there is a stark difference between them in information transparency and political accountability. In democratic countries, media and social media that are independent from the political system disclose the factual developments of the intrusive event. And they are also dedicated to expose policy failures in relation to how government agencies and agents deal with the event. Society, through learning such information, can hold political leaders accountable in elections. In authoritarian regimes, however, dealing with intrusive events has one priority, which is to restore authoritarian rule as quickly as possible.

In authoritarian regimes, information about intrusive events largely flows in a unidirectional, top-down manner. Society, for its part, tends to accept whatever information it is given, provided it is benign and conforms more or less to its own experience. On the other hand, because society knows that the regime has total control of the information, it can also be

skeptical of it. When the skepticism becomes serious and widespread, the regime can and sometimes does make adjustments. In China at least, the iron hand shows remarkable flexibility—this is how it maintains power!

To the high generalization above, I now offer a special case—an experiment—to understand in detail society's confidence in the information autocrats provide in the face of an intrusive event. That event is a terrorist attack on a densely populated area where eye-witnesses are likely to be the first on the scene, report what they find through the social media, and thus somewhat curtail the regime's ability to manipulate information for its own purpose. I focus on information manipulation of this intrusive event in cyberspace because it is there that the state media break the news and netizens first respond with their opinion. Social media on smartphones have in recent years become the gateway of information for Chinese society. Consider the fact that China's netizens, a population of 668 million by mid-2015, 89 percent of whom have access to the internet on smartphones. Empowered by such access, netizens are able to gain information of a sudden, highly stressful event from other netizens and sometimes even from Western media, in real-time, regardless of their whereabouts, before autocrats can apply cyber-control. Given the enormous gain in communicative power, isn't it plausible that netizens, triggered by an intrusive event, may engage in collective rebellion in cyberspace that challenges authoritarian rule?

Future Research Agenda

Information, Repression, and Rebellion. I have concentrated so far on information contestation between autocrats and netizens in cyberspace. Beyond the information that both protagonists possess are the actions that follow. Should autocrats use force, threat, or political coercion against civil society, a consequence will be a shift in their political organizations and institutions. Once this happens, there is always the prospect of further political shifts and

challenges ahead. As for disgruntled netizens, they show their displeasure and anger by making organized demands. To better understand the actions undertaken by both autocrats and netizens, I turn to sudden events, which may either be internal such as a purge in leadership or a political coup, or external such as a terrorist attack or a natural disaster. All such sudden events are destined to shake up, if only temporarily, unquestioned authoritarian rule. Also to be considered is how autocrats institutionalize procedures in cyberspace to resolve their contestations with netizens. Such contestations are bound to increase as the number of young, mobile, and educated netizens, already a large number, continues to increase. One adaptation on the part of Chinese autocrats is to introduce a tribunal system on social media to handle political disputes and to give such handling the semblance of "due process." Another example is to introduce auctions online of the goods and properties confiscated in the working of their judiciary system. Will such institutional adjustments and innovations be able to sustain authoritarian rule? I will examine this question in my future research.

Communication. Internet, social media, and smartphone have changed the means of communication. When scholars overwhelmingly focus on the 'media' perspective of social media—a form of communication that is faster in reporting and bigger in response—they neglect the 'social' perspective of social media, with 'social' being a large number of users that only contribute a small proportion of posts on social media. Social media electronically conveyed allow atomized individuals who share a physical space to recognize, instantly, one another's existence and so produce a social network (or community) by digital means that once depended on slow face-to-face contact. In today's world, electronic bonds can grow very large quickly. As my study shows, when confronted with authoritarian control, communities resort to disguising their communication, which is clearly a sign of distrust; on the other hand, they may disguise

their belief by inflating their support. As new technologies such as virtual reality (e.g. 360-degree picture) and location-based services (e.g. Uber or Pokemon Go) further change communication, their impact on politics, society, and geography is bound to increase, thus raising new questions.

Computational Social Science. New communication technologies offer us unprecedented amount of data and metadata that come with a much finer resolution over time and space than what were previously available to social science. As my dissertation shows, data of different kinds can be combined to produce a method that makes it possible to map communities in the less developed parts of the world and in authoritarian countries such as China that are reluctant to part with cartographic information. Of more direct importance to socio-political science is that the new electronically recorded data tell us something of what people think and feel at critical times and places. In this regard, they differ from data obtained through survey and interview methods that necessarily post-date the critical event by months and years.

Geography. The meaning of place is changing in this increasingly digital era. Do mobility, communication, and information end geography? My dissertation points to the opposite: places are the focal points of political contentions. As for political protesters, they congregate in public squares, government buildings, and embassies, places of high visibility and resonance that seem able to amplify their message such that, for example, an imposing, colonnaded government building in the background personifies that stony authority the protesters object to in their governments. Worse for the government is that if it uses force, that force is only made more brutal when seen against the backdrop of a stone-faced building. Not only do people speak, places can do so too: they too are the media. In China, Tiananmen Square outstandingly illustrates this point. A vast square with a history of authoritarian power, it has many uses,

including the display of military might. On most days, however, it is where the children play under the solicitous eye of their parents and where grown-ups pose for their selfies under the solicitous eye of Chairman Mao.

Concluding Statement

As I note above, I believe that my contribution is "a view from above," which I also take to mean the perspective of cyberspace. Both are abstract, relative to detailed interviews and surveys. At the same time, I find myself moving in the junction of political science, communication, computational social science, and geography—of places as media in political communication, which is the direction of the concrete. Perhaps by moving in both directions I can contribute most to our understanding of the sociopolitical reality.

MAPPING COMMUNITIES IN RAPIDLY DEVELOPING COUNTRIES BY USING MEDIUM-RESOLUTION REMOTELY SENSED IMAGERY AND GEOTAGGED INFORMATION

1. Introduction

Geographers are keenly interested in mapping communities, for their size, character, and change can shape economics and society in rapidly developing countries. Communities extends across cities and villages, and their transformation, expansion and contraction reflects sociological, political, and economic forces at work. These forces, whether conforming to governmental policies or not, have social, economic, and political consequences. All researchers desire communal information that is accurate and can be absorbed, if possible, almost at a glance. This is information presented in graphic or cartographic form. How to achieve such desirable presentation depends on the availability and quality of data and on the method used. This is not a serious problem in developed countries where rich sources of data are available to the researcher. In developing countries, however, the data are less ample and less available (Montgomery 2008), and this includes China, even though China has a real census of communities as early as in 2AD (Bielenstein 1954). Yet, it is precisely in these countries that communities have transformed the most in recent decades (Angel et al. 2011). Called for, then, is the optimal use of the data that are available and a new method capable of doing them justice. First, however, I state briefly the limitations of the remote sensing technique. I follow this with another brief statement about what big data and my new method can do to overcome them.

Mapping communities requires big data. In developed countries, researchers and urban planners make arduous work in mapping communities with surveys and public records. In developing countries, documentation and survey are sparsely and irregularly distributed across space and time and thus have little use of mapping the communities with the satisfactory degree

of consistency and accuracy. My method integrates the recent advances in land cover mapping, large-scale cyber-Geographic Information Systems (GIS) data, and volunteered geographic information (VGI).

There has been increasing sophistication in the mapping of land cover change by using the technique of remote sensing such that land cover categories—epitomized by forests, water, built-up areas—can now be detected at a level above 75 percent in accuracy across methods (Mountrakis, Im, and Ogole 2011). Given the high level of accuracy, delimiting and analyzing land cover change within each major category should be achievable. Urban land use change, however, is not the same as urban land cover change (Lambin et al. 2001). Land cover is defined as coverage of land surface, without specific reference to human use. In contrast, land use is the human use of land, which is especially prominent and complex in urban environments. The divergence poses two challenges: the same land cover has different land uses; the same land use has different land covers. To give an example of the first challenge, an airport is identical to dense commercial and residential area in terms of land cover, but clearly differs in regard to land use. As for the second challenge, upscale communities have land cover, including built-up areas, grass, and water in surrounding lawns and gardens, that is different from urban slums immersed by built-up areas, though both would be classified as residential use. This conundrum leaves an impasse for applying remote sensing to urban land cover and land use studies.

The arrival of volunteered geographic information opens possibilities to disentangle urban complexities (Goodchild 2007; Batty 2013). One complexity which I specifically wish to address is to identify specific land use that, under land cover, appears as built-up areas. The points of interest (POIs) that are mapped by online mapping services provide locality-dependent information (Arsanjani et al. 2013). This information, though always incomplete and sometimes

inaccurate, contains details of current land use that is unprecedented. In addition, the rise of smart phones and social media spurs conversations across time and space, some with geotagged information, which provides a glimpse into the spatiotemporal pattern of a sample of urban residents (Song et al., 2010). Such a glimpse may in turn help us to infer land use.

With the opening up of these new datasets, it is now possible to advance from mapping urban land cover change to mapping urban land use change. To achieve the latter, I propose a new approach that integrates urban land cover change classification with other data such as POIs, road networks, social media data, and types of housing. Such an integrative approach far transcends a simple reliance on the use of remote sensing images, however improved.

My approach raises and addresses such questions as: how to classify land cover change in fine detail so that land lots that share the same land use can be delimited; how to map major land use categories using POIs; how to map land use sub-categories such as gated communities, ordinary communities, and urban slums within residential use; and how to assess accuracy. My findings are presented in the following order: “Study Area and Datasets” introduce the geographical extent of their potential application and the newly available data; “Method” provides the detailed steps in its implementation; “Results in cartographic presentation” are then given, followed by “Assessments” that address the limitations and potential of my method.

2. Study Area and Datasets

2.1 Study area

I select Kunming, China to demonstrate my approach because it is an example of an urbanized area that has seen complicated, dynamic changes in land use in the past few decades. Geographically, Kunming is seated on a mountainous plateau in subtropical China (Figure 1).

[Figure 1 about here]

Kunming has been a resource poor, partly due to its location in the rugged landscape of the Yungui plateau. It has a small amount of agriculture near the core city on the edge of Dianchi Lake, but is surrounded primarily by forest. Historically, it has been a significant military base and a trade center located on an ancient route for tea. Today, Kunming is a provincial capital of southwestern China, with a population of 6.53 million at the end of 2013 (Kunming Statistical Bureau 2014). The economic reforms in the late 1970s and in the 1980s triggered rapid urbanization and construction. Over the past three decades, Kunming's population doubled in an economy that soared (Kunming Statistical Bureau 2014).

2.2 Remotely sensed imagery

Medium-resolution remotely sensed imagery—Landsat Thematic Mapper (TM) /Enhanced Thematic Mapper Plus (ETM+)—is used to classify land cover changes with special regard to consistency, spatiality, and temporality. Compared to Light Detection And Ranging (LiDAR) data (<1m resolution as of small footprint), high-resolution images (<4m, e.g. Quickbird), and coarse resolution images (e.g. Moderate-resolution Imaging Spectroradiometer, or MODIS), Landsat data remain the best option for tradeoffs in strengths and weakness. Strengths include: (1) images with the same spatial resolution since the 1980s, (2) consistent timespan coverage for my study, (3) temporal intervals between data, which range from 16 days to 6 months, shorter than the construction period of buildings, and (4) open access of Landsat archives since 2008. These strengths enable us to capture urban construction at different times of groundbreaking. Moreover, the pixel size of TM/ETM+, 30 meters, can capture construction from single-family houses (usually bigger than 15m in length and 15m in width) through apartment buildings (usually bigger than 100m at length and 30m at width) to multi-building complex (acres big). I collected all cloud-free, terrain-corrected Landsat data from 1988 to 2013, as reported in Table 1.

[Table 1 about here]

Besides medium-resolution remotely sensed imagery, I also collected Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Digital Elevation Model (GDEM), the digital elevation model (DEM) at 30 meters, as an auxiliary dataset. Its topology helps select training samples semi-automatically for forest land cover at certain slope, aspect, and altitude, as guaranteed by Chinese Grain for Green policies (Xu et al. 2010)[1].

2.3 Cyber-GIS and VGI data

I also used a variety of geographic data that only emerged in recent years but have since exploded in abundance. The data, including roads and POIs, come from three major sources: OpenStreetMap, Google Maps, and Tencent Map (Table 2)[2]. Compared to the voluntarily contributed OpenStreetMap and closely surveilled Google Maps, Tencent Map is competitive in that it has been expanding its POIs and street view images, and even provides turn-by-turn Global Positioning System (GPS) service (Millward 2014). As for roads, the datasets from the three sources are rather complementary, after geometrically correcting them with Landsat images, yielding a detailed network that delineates not only boulevards but alleyways and small street blocks in the downtown area (>1,800 m^2). As for POIs, Tencent Map is superior to the other two sources in terms of completeness. In addition, categorically, every POI in it is labelled with three levels, with increasing categories of land use at each level. For instance, a Cantonese restaurant is labelled “Food” at level 1, “Restaurant” at level 2, and “Cantonese cuisine” at level

[1] Grain for Green (*tuigen huanlin huancao*, restore the farmland to forest and grassland) policies forbid farming and herding practices, as well as construction, on slopes steeper than a certain degree. As an alternative, the policies subsidized reforestation and grassland restoration, according to the natural environment.
[2] I originally used SOSO Map services, which was subsequently acquired by Tencent and incorporated to Tencent/QQ Map. Therefore, I use Tencent Map in this article.

3 (Tencent 2014). Given the detailed information provided and to avoid redundancy, I only use POIs from Tencent Map.

[Table 2 about here]

Another variety of geographic data derives from geotagged social media, specifically, from SINA Weibo, the Chinese equivalent of Twitter, which I used through its API from January 2014 to May 2014. The location of the post comes with geographical coordinates measured by cellular or Wi-Fi triangulation of its sending device. Its error, by estimation, ranges from 2 to 10 meters, depending on conditions such as physical barriers, the location of the cellular towers, or Wi-Fi router types, for example (Zandbergen and Barbeau 2011). Given that the resolution of Landsat image is 30 meters, the horizontal error is negligible for the purpose of my research.

In addition to the variety of geographic data noted above, I also take advantage of aspatial data that are informative to my approach. I mapped these socioeconomic data to further differentiate residential uses. My method entails collecting the prices of mobile devices as listed on Chinese e-commerce websites (e.g. Taobao.com and JD.com, Chinese equivalent of Amazon.com) on May 21, 2014, as well as a variety of information about housing projects such as building stories, number of bedrooms and living rooms in a sample of units, recorded sold price, and sale price per square meters of old and new units, either residential or office or commercial, as listed on Chinese real estate sale and rental websites (e.g. Fang.com, Chinese equivalent of Realtor.com) on the same date. For the purpose of evaluation, land use planning maps of Kunming, made at different times, were consulted and georeferenced with Landsat images. Also used in this study are Google Earth historical images and street views from Tencent Map.

3. Method

With all data collected and preprocessed, I take three interwoven steps to map urban land use change. The first step is to classify the change in land cover. The second step is to map current land use into such major categories as industrial, commercial, office, transportation, recreational, and residential. Each major category has in fact mixed land uses; for example, the commercial use category may include residences. For the purpose of this article, which is to demonstrate the power of a method, I choose for detailed consideration one major category—residential, which itself may be subdivided into gated communities, ordinary communities, and urban slums. The three steps of my method are presented as a pictograph, with each applet referring to each step (Figure 2). In the rest of this section, I develop essential background, hypotheses supported by the literature, instructions, and analysis that elaborate on each step.

[Figure 2 about here]

3.1 Land cover change classification

First step. I classify the land cover change at minimal year intervals by using Landsat images, a procedure that requires us to use supervised classification algorithm to map spectral distinctions across major land cover types. If a medium-resolution pixel (900 m^2) contains mixed objects such as built-up areas, grass, and soil, the built-up land cover is defined provisionally as wherever over 50 percent of a pixel is made of manufactured surfaces such as concrete or asphalt. To further improve the classification, I include in it land cover types other than built-up areas, such as water, forests, grass, and exposed soil.

For each land cover class, a large number of pixels that represent each targeted land cover type are required as training samples, the manual selection of which is a time-consuming process. To expedite the process, I extract ancillary datasets to stratify the selection by using Normalized Difference Vegetation Index (NDVI) in two periods, 1988-1993 and 2010-2013, to

filter out areas that stay in both periods at overall high value as vegetated land and low value as non-vegetated land (Galford et al. 2008). To select samples for urban land, I mask out areas enclosed by small street blocks and select more than 213 individual pixels from non-vegetated land. Out of these small street blocks, the non-vegetated areas are my candidate sample pools (twenty-four pixels) for exposed soil. The vegetated areas are my candidate sample pools for forest, which reveals itself in having relatively low seasonal variations of NDVI in each period. As for areas where NDVI variations are relatively high, they are my candidate for agricultural land (Galford et al. 2008). Using DEM, these distinctions can be further refined, as follows. Slopes that range from 17 percent to 28 percent (15 degree to 25 degree) and lie at an altitude below the tree lines are mandated by Grain for Green policy to be forest (Sherman et al. 2008). Given this policy and its enforcement, I am justified in selecting a sample of 193 pixels from such slopes for stable forest. As for slopes gentler than those covered by forest, I select a sample of forty-three pixels for stable grass. Areas that have extremely low values in NDVI (areas that are more or less flat) I select a sample of seventy pixels for water body (Huang et al. 2010).

With all training samples selected, I am better positioned to classify images into a number of small, piecemeal land cover changes at very fine spatial and temporal scale. I use Support Vector Machine to classify each image (Huang and Townshend 2002; Wu, Lin, and Weng 2004). I then check the consistency of the result of each image with previous and posterior time points to eliminate error caused by cloud or scanner liner corrector-off (SLC) strips. The next step is to integrate the results for one year (or minimal years if data are limited), a procedure that requires us to designate the land cover as built-up or non-built-up, depending on the most frequent land cover type among the results in a particular year. For instance, after integration, if a parcel of land is consistently non-built-up in January, March, and May, but converts to built-up

in October, with all results correctly classified, it is classified as non-built-up for that year; in the following year, the conversion from non-built-up to built-up is recorded. Lastly, to keep that larger parcel (>= 4500 m^2) intact, single pixels in adjacent years are annexed.

3.2 Major categories of land use classification

Second step. With urban land cover successfully classified, I can now lay the foundation for delimiting land parcels, within which land cover changed from non-built-up to built-up during the same year, as the land lot that may share the same land use. The method, however, cannot discriminate adjacent land parcels that have different land uses in two conditions: (1) these parcels were built in the same year; (2) these parcels were in the prominently long-term urban area, which mostly consists of built-up areas during my research period. To overcome these difficulties, I make use of the current road networks, which are especially dense in the core area, for two reasons. One is its relative pattern persistency and other is that it dissects the urban core area into blocks and further splits the land parcels built in the same year along two sides of the road. At the end of all these classificatory procedures, I am in a position to consider land parcels—either the small pieces of land developed in the same year or the blocks in the downtown area with dense road network—as the basic spatial elements of land parcels that share the same land use.

I now proceed to identify current land use. Rather than delimiting numerous land use types, however, I collapse them into major categories as these are defined in urban planning (Table 3). A reason for using urban planning's definitions is that they reveal the land's essential economic and social functions. As an example, take museums and parks, which, though they have hugely different functions under land cover classification, serve the same function under land use classification, as both provide amenities for urban residents and tourists. As I noted in Section 1, among major land use categories that are mixed, I further take residential as an

example to identify land use sub-categories, particularly in regard to their profoundly different economic and social functions. Three residential sub-categories are recognized for the purpose of this study: gated communities, ordinary communities, and urban slums. Gated communities are mostly walled real estate projects. They take the form of either multiple high-rise buildings or small condo apartment buildings surrounded by a landscape of aesthetic elements such as fountains and meandering creeks (Zhang 2007, 109-115). Ordinary communities are apartment buildings of six to eight stories, built in Soviet-era style and compactly arranged. They used to cater mostly to workers who belonged to work units, now tenement occupants. Urban slums are small houses of one to three stories, confined in small urban spaces, and of substandard construction (UN-Habitat, 2004).

[Table 3 about here]

With the sub-categories of urban residential use thus defined, how are they to be separated from other categories of urban land use? Moreover, how are the residential categories themselves to be distinguished from one another? To address these questions, I first need to identify land use at the scale of a land parcel. This calls for the assistance of POIs. Often a land parcel is overlaid with (or surrounded by) a number of POIs, each carrying descriptions of land use in the same three-layer structure. To identify a land parcel with informative POIs, I summarize POIs in structures that correspond to these categories.

Since each POI indicates a different amount of land use information on the ground, I categorize them based on spatial size. For instance, a POI of a real estate community virtually indicates a large area that contains a number of residential buildings, while that of a hair salon only indicates a tiny store among other stores on the first floor in a building. Based on the spatial size of POIs, I categorize them as building complexes, individual buildings, and sub-buildings.

After categorizing, only a handful of POIs fall into the category of building complex, such as factories, real estate communities, airports, as their architectural structure may extend to acres large or even occupy a street block. Other than this, dozens of standalone buildings, such as museums, libraries, and gymnasiums, fit into the category of individual building, as they exhibit as a building or connected construction. The majority of POIs, many of which often co-occupy a single building, belong to the sub-building category, such as companies, restaurants, hair salons, and so forth.

As categorized, POIs constitute a hierarchy of combinations on each land parcel that are informative about the major land use categories. Drawing upon instances from urban planning maps, land use maps, and fieldwork, I summarize rules for each type of land use, in configuration and composition with various POIs, and illustrate them in Figure 3. I selectively use POIs that are positioned on or near the land parcel in order to determine its land use, instead of using all POIs in equal weight. A POI of multi-building level trumps POIs at a lower level in terms of defining the land use of land parcels underneath it or nearby. For instance, a POI of real estate community surrounded by a number of restaurants and shops still posits a great likelihood for the land parcel to be of residential use. With these rules set up, I am able to classify targeted land uses.

[Figure 3 about here]

3.3 Residential category and its sub-categories

Third step. As major land use categories have been identified, the use that represents each land parcel in fact only stands for the primary use of land in mix of land uses. Take an average office building as an example. It may host cafeterias and restaurants, besides a large number of offices. To further differentiate the sub-categories of land uses, I use residential use as an example. Unlike other land uses, residential land use is especially underrepresented for three

major reasons. First, the residential community may consist of uncertain numbers of land parcels that were developed in adjacent years. Yet not all of them spatially intersect with its single POI. Second, the POIs of residential use are often placed next to streets (for convenience to provide the turn-by-turn GPS orientation), instead of at the center of land parcel(s). The land parcels away from the streets are underrepresented. Third, for some ordinary communities and almost all urban slums, there are no POIs of residential use to represent them. For instance, the housing units and dormitories in schools, hospitals, and state-own enterprises are left out, as the POIs nearby or on top simply represent schools, hospitals, state-owned enterprises, and other uses. In the same way, the residential part of small restaurants, convenience stores, and hair salons in urban slums is neglected by the fray of other land uses.

To better classify the residential use that has been underrepresented, I incorporate social media data and use its temporal patterns over each land parcel. Hypothetically, the temporal pattern of posts at home differs from that at work or other places in three ways. First, users at their residence are likely to send more posts off work hours—in early morning and late at night (specifically 11pm-2am) than during work hours (specifically 9am-12pm) on weekdays. Second, they are likely to send more from 9am to 5pm during weekends than during weekdays. Third, they are likely to send more during holidays and special occasions when they are supposed to stay at home for family gatherings (e.g., the Eve of Spring Festival). Based on these criteria, I designate land parcels as residential to the extent that the temporal pattern of posts satisfies two of the three ways noted above.

Once completed, I further classify residential use into three sub-categories, namely, gated communities, ordinary communities, and urban slums (Table 3). Gated communities are characteristically grand in terms of building height and unit size and exclusive in term of price

and surrounding services. Ordinary communities are characteristically six to eight stories tall and built in the Soviet style, with the lower levels used as stores and restaurants. Urban slums are characteristically substandard dwellings, one to three stories tall with street-level shops and food stalls.

In order to locate and map gated communities, I use detailed and specific information, which includes the following: (1) the real estate listing record, which supplies information on basic construction type—whether the buildings are villas or high-rise (>10 stories), apartment unit price, number of bedrooms and living rooms; (2) POIs, which supplement the information by indicating the type and number of surrounding services (for instance, a large number of services suggest commercial development, not gated communities); and (3) the price of sending mobile devices, which enables us to gauge the wealth of the residents.

I locate and map ordinary communities as those residences that remain after I have mapped gated communities through POI. A special challenge of China before my period of research is that POIs that designate schools, hospitals, and state-owned enterprises may in fact also be residences. But whether they are residences or not can be tested by using social media. Finally, as the urban slums are never designated by POIs but are recognizable by the posting pattern, I classify the remaining residential uses as urban slums.

3.4 Evaluation

I evaluate the spatial and temporal accuracy of my method by first separating the urban core area from the peri-urban area. I take this step because the peri-urban area, which refers to the area that has been converted from non-built-up areas to built-up areas after 1988, the year my research starts, can be monitored accurately by remote sensing. In the peri-urban area, I can take advantage of Google Earth Historical Images in very high spatial resolution, which, however, are available only from 1999 to 2014 in my study area). My procedure is to select points for

evaluation in a three-year interval consistent across the peri-urban area. Specifically, I select five points for 2001-2003, twelve for 2004-2006, twenty-eight for 2007-2009, and thirty for 2010-2012. In contrast, the core area where the land cover change is a matter of urban renewal, that is, from built-up to built-up area, the changes that occur cannot be tracked by Landsat images, and I am left to depend on dense road networks and online real-estate trading record. To test for accuracy, I generate fifty-three random points in the urban core and compare them with available Google Earth Historical Images (Potere 2008).

I also evaluate the accuracy of urban land use classification in both major categories and residential sub-categories. For accuracy, I employ Tencent street view, Google Earth very-high resolution images, and Kunming planning maps. Given that the street view only provides the look of the street in 2012, that Google Earth images provide only a top view but not land use, and that planning maps are limited, I can only manage to evaluate accuracy by carefully checking whether the land use in my map is consistent with all three visual sources just noted. To do so, I proportionally selected a number of points of each land use type in my map: eighty-nine points in commercial use, sixty-one in office, twenty-six in manufacturing, forty-five in transportation, and twenty-six in recreational use. For the sub-category of residential use, I selected eighty-seven points in gated communities, eighty-nine in ordinary communities, and eighty-three in urban slums. I then assessed their accuracy by comparing them, again, with the three visual sources.

4. Results in visual representation

The method I have employed enables us to produce a number of results, which I present in tables, figures, and maps. These are at the core of my findings. They show, in essence, urban land use change in a spectrum of depth and finesse, ranging from the surface of land regardless

of specific use, through land use in its major categories, to land use in a specific category, which, in my case, is residential.

Whereas tables and figures speak for themselves, maps call for further comment. They display three themes: (1) expansion of built-up areas at the periphery shown at a fine spatial and temporal granularity, which in my research is at 30m pixel-size and at one-year or minimal years intervals. In detail and precision, what I have done improves significantly on the mere conversion of non-built-up to built-up land cover in the urban periphery (peri-urban) and of addressing only the one-time conversion of land cover from one to the other; (2) current land use in major categories, including manufacturing, office, commercial, transportation, recreation, storage and farm houses, and residential; and (3) one major category—the residential and its sub-divisions: gated communities, ordinary communities, and urban slums. These three themes are further elaborated in the following three sections.

4.1 Expansion and renewal of built-up areas

Expansion in the peri-urban area, which entails change of surface from non-built-up to built-up, is mapped by using remotely sensed imagery. I took spatiality, temporality, and accuracy into account through an integrative use of phenology, topography, and road networks, the training sample of which is selected in a timely and consistent fashion. A posteriori consistency check is then applied to close gaps caused by cloud and SLC, and this is possible because the land cover change from non-urban to urban in China is largely unidirectional (Schneider, Chang, and Paulsen 2012). Pre- and post-classification processing enables different classifiers to capture the majority of land cover change, with the accuracy of annual change Support Vector Machine (SVM) achieving 78 percent with ten times cross-validation in the peri-urban area. Cumulatively, the accuracy by 2013, the end of my research period, reaches above 90 percent across classifiers.

As for the urban core, where the change is from built-up areas to built-up areas, I took further steps by annexing single pixels to adjacent continuous pixels in the same urban block if the changes occurred in nearby years. The annexation is based on the rationale that these single pixels register mostly construction extensions of major buildings. Chinese cities, being heavily planned, are unlikely to allow single-family houses the size of single pixels to be constructed separately. I enlisted POIs and the built-years of real estate to understand the approximate year when the renewal occurred (integrative approach in Figure 4). By investigating an additional fifty-three randomly generated evaluation pixels that fall under the built-up areas of the urban core during the research period, my approach correctly classified forty-eight pixels, including nine in which urban renewal occurred.

[Figure 4 about here]

To visually compare results from SVM and the integrative approach I just noted, I select three typical areas in an urban corridor that stretches from the downtown area to the suburban and rural area. These are: an urban core that has undergone periodic renewal (Figure 4A); a peri-urban area that has undergone rapid expansion (Figure 4B); and a planned new town that is located a dozen kilometers from downtown (Figure 4C). My results show that after time stacking and post-classification adjustments, a majority of built-up changes over the years has been effectively captured. SVM produced consistent results across space, which suggests that it is able to exploit the feature distinctions across training samples that have been differentiated by NDVI and so efficiently offset the cloud and SLC errors in the post-classification process.

4.2 Land use in major categories

Having considered change in built-up areas, I now turn my attention to how non-built-up areas developed into built-up ones and address the classification of major land use categories. By pruning POIs, connecting road networks, and assembling social media data in an integrative

approach, I am able to obtain results in land use classification of acceptable accuracy. Overall, it reaches 73.6 percent. Variation is, however, large from one category to the other (Table 4). Thus, in commercial, office, transportation, storage and farm houses, and residential categories, accuracy reaches levels of 86.6 percent, 67.4 percent, 65.9 percent, 69.2 percent, and 83.6 percent respectively. This is satisfactory, but in recreation and manufacturing, it is lower.

[Table 4 about here]

To visually inspect the result, I show a map that covers both the downtown and suburban areas of Kunming in major land use categories (Figure 5A). It is quickly obvious that the downtown area is primarily for residential use, with, however, such other uses as upscale commercial, office, and a recreational area reserved for structures of historical significance, for the city's draw—its function as a hub—goes beyond mere housing. Outward from downtown, land uses are mixed and include residences and industries that have had to move out with urban expansion, newly-built residences for new residents, as well as, old stores and factories that have somehow escaped removal, new commerce and industries. Still farther out from the city core is the suburban area where land uses are even more mixed and include villages, more or less isolated farm houses, new communities, offices and industries that seek not only cheaper land but also escape from the sort of close governmental regulation that is likely downtown.

[Figure 5 about here]

4.3 Residential sub-categories

Residential, a major category in land classification, is divided into three sub-categories (Figure 5B). I do so, as I noted above, because of my interest in the socio-economic structure of a Chinese city and its possible political implications. But I take this step also because it shows how the finer delineation of a major category—the residential—becomes possible with the use of real estate listing data and social media data. The result of my method is shown in Figure 6.

Gated communities are classified at an accuracy of 72.7 percent (the yellow fan-shaped area), ordinary communities are classified at an accuracy of 72.4 percent (the orange fan-shaped area), and urban slums are classified at an accuracy of 88.7 percent (the brown fan-shaped area).

[Figure 6 about here]

A visual inspection of Figure 5B shows how these communities are distributed. In the urban core, where a large number of POIs are most indicative of land use and where social media are most abundant and reflective, gated and ordinary communities are clearly distinguishable even though they occupy already developed space. In the suburban area, where built-up patches are interrupted by such natural features as lakes and green space, gated and ordinary communities are even more clearly bounded and so distinguishable from one another. Urban slums, by contrast, are fragmented and shapeless. They do not fit into zones, and appear to scavenge urban spaces that are either undesirable or are otherwise occupied.

5. Assessments

My approach integrates a variety of geographical data and techniques that have become available in the last three to five years. These include not only datasets, both new and the newly accessible, but also knowledge provided by various disciplines and, importantly, a willingness to reason and interpret, based on a considered sense of likelihood or probability that is itself based on a habitual use of and exposure to statistics. What have I achieved? What are the limitations of my method? What is its potential? I conclude with a number of brief assessments, beginning with a few achievements not specifically addressed earlier.

5.1 Change in built-up areas horizontally and vertically

To tackle urban land use change, my method starts from non-built-up areas to built-up areas. The process is called urban expansion, and it is usually treated as occurring along a horizontal plane. As such, the expansion is easily and accurately recorded by techniques of

remote sensing. However, urban expansion is also vertical, which is less noted in land cover change studies. My method opens up the possibility of understanding this kind of land cover change—the change from built-up areas to built-up areas, with the assistance of dense road networks, POIs, and real estate listing records. Relying only on the publicly accessible data, my method proves to be effective under three circumstances: that the change be solely caused by the construction of gated communities and newly paved roads that circulate these communities; that the change be led by the construction of residential communities together with services—such as stores and food stalls—that were built at the same time; and/or that the built-up areas remain unchanged during the research period.

5.2 The rise of multiple or big data

Observing the change in built-up areas gives us little understanding of such urban complexities as where urbanites of different socio-economic classes live and work, and whether or how they commute. My method sheds a new light on these complexities thanks to the multiple data (or big data) that have become available in recent years. The multiple sources of data I used in this research significantly increased my analytical powers. They are especially necessary to developing countries where data are restricted or are not sufficiently reliable. For instance, the Google very-high-resolution image provides a free-of-charge and reliable reference for assessing my result. In addition, data, even if not new, may provide a complementary means for research. For instance, in China, data on residents at the community level has never been available in any officially published census or yearbook, and yet, for a certain class of population, social media allow us to peek into the spatiotemporal pattern of their lives. As volunteered geographic information and location based services keep expanding, I am in an ever better position to understand social and environmental reality even before I set foot in critical places for field observation.

5.3 Limitations

The multiple sources of data I have use in this research has significantly increased my analytical powers and yielded information that is not otherwise available or accessible, this being especially the case in developing countries. I am, of course, aware of the limitations—data, algorithmic, and analytical—in my method. One data limitation is incompleteness, such as in cloud-shadowed remotely sensed imagery, underrepresented POIs, and communities that are not documented in real estate records. Another data limitation is one of error, which may be numerical, incorrectly placed, or mislabeled. As for algorithmic limitation, one that detracts from my study is that I am unable to consider the construction period, which may take three to five years in gated communities, for example, and which varies from project to project. Another algorithmic limitation is that, for lack of information, I cannot proceed to find what the land use was prior to urban renewal. As for analytical uncertainty, I encounter it most forcefully at the urban periphery where boundaries between a wide variety of uses—office, commercial, slum, and rural houses—are vague.

5.4 Potential

Integrative mapping of complex and changing urban phenomena lies at the core of my article. The word “integrative” is important because it is unlikely that a single method or technique, however sophisticated, can fully capture urban reality of the sort that addresses socioeconomic and political issues. The word “mapping”, which is to say, cartographic representation, is important because it enables the researcher to see, at a glance, the changing size, location, and configuration of the type of urban phenomena the researcher seeks to understand. The understanding so sought may be relatively straightforward; for example, the infringement of expanding urban area on agricultural land, and it can be complex such as seeing

how different types of urban community serve as indicators of wealth, social status, and, possibly, political stability.

In developed societies, it may be that sufficient data of high quality have been collected and are, moreover, available to the researcher so that the integrative approach I propose is not essential. In developing societies, this is not the case. It is there that ingenuity needs to be exercised to extract information from multiple sources, including those that may be unique to a people. The information so obtained may then be used to supplement or confirm the more standard methods of mapping. My article shows how this can be done, using Kunming, a city of modest size (3.5 million) in rapidly developing China as an example.

REFERENCES

Angel, S., J. Parent, D. L. Civco, A.Blei, and D. Potere. 2011. The dimensions of global urban expansion: Estimates and projections for all countries, 2000–2050. *Progress in Planning* 75(2): 53-107.

Arsanjani, J. J., M. Helbich, M. Bakillah, J. Hagenauer, and A. Zipf. 2013. Toward mapping land-use patterns from volunteered geographic information. *International Journal of Geographical Information Science* 27(12):2264-2278.

Batty, M. 2013. *The New Science of Cities*. Cambridge: MIT Press.

Bielenstein, H. 1954. The restoration of the Han dynasty: with prolegonmena on the historiography of the Hou Han Shu. *Bulletin of the Museum of Far Eastern Antiquities* 26:1-210

Galford, G. L., J. F. Mustard, J. Melillo, A. Gendrin, C. C. Cerri, and C. E. Cerri. 2008. Wavelet analysis of MODIS time series to detect expansion and intensification of row-crop agriculture in Brazil. *Remote Sensing of Environment* 112(2):576-587.

Goodchild, M. F. 2007. Citizens as sensors: the world of volunteered geography. *GeoJournal* 69(4): 211-221.

Goodman, D. S. 2004. The campaign to “Open up the West”: national, provincial-level and local perspectives. *The China Quarterly* 178:317-334.

Huang, C., L. S. Davis, and J. R. G. Townshend. 2002. An assessment of support vector machines for land cover classification. *International Journal of Remote Sensing* 23(4):725-749.

Huang, C., S. N. Goward, J. G. Masek, N. Thomas, Z. Zhu, and J. E. Vogelmann. 2010. An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks. *Remote Sensing of Environment* 114(1):183-198.

Huang, Q., 2016. Mining online footprints to predict user’s next location. *International Journal of Geographical Information Science*: 1-19.

Kunming Statistical Bureau. 2014. *Kunming Statistical Yearbook 2013*. Beijing: China Statistics Publishing House.

Lai, H. H. 2002. China's western development program: Its rationale, implementation, and prospects. *Modern China* 28:432-466.

Lambin, E. F., B. L. Turner, H. J. Geist, S. B. Agbola, A. Angelsen, J. W. Bruce, O. T. Coomes et al. 2001. The causes of land-use and land-cover change: moving beyond the myths. *Global environmental change*, 11(4), 261-269.

Millward, S. 2014. Tencent hits the road (and goes off the beaten path) to create China's biggest 'street view'. Accessed January 25, 2015. http://www.techinasia.com/Tencent-maps-covers-china-street-view/

Montgomery, M. R. (2008). The urban transformation of the developing world. *Science*, 319(5864): 761-764.

Mountrakis, G., J. Im, and C. Ogole. 2011. Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing* 66(3):247-259.

Potere, D. 2008. Horizontal positional accuracy of Google Earth's high-resolution imagery archive. *Sensors*, 8(12), 7973-7981.

Schneider, A., C. Chang, and K. Paulsen. 2015. The changing spatial form of cities in Western China. *Landscape and Urban Planning*, 135, 40-61.

Seto, K. C., and J. M. Shepherd. 2009. Global urban land-use trends and climate impacts. *Current Opinion in Environmental Sustainability* 1(1):89-95.

Seto, K. C., B. Güneralp, and L. R. Hutyra. 2012. Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. *Proceedings of the National Academy of Sciences* 109(40):16083-16088.

Sherman, R., R. Mullen, L. Haomin, F. Zhendong, and W. Yi. 2008. Spatial patterns of plant diversity and communities in Alpine ecosystems of the Hengduan Mountains, northwest Yunnan, China. *Journal of Plant Ecology* 1(2):117-136.

Song, C., T. Koren, P. Wang, and A. L. Barabási. 2010. Modelling the scaling properties of human mobility. *Nature Physics* 6(10):818-823.

Tencent Corporation, 2014. The developer's guidance of Tencent map. [in Chinese] Accessed January 25, 2015. http://open.map.qq.com/doc/guide-map.html

UN-Habitat. 2004. The challenge of slums: global report on human settlements 2003. *Management of Environmental Quality: An International Journal* 15:337-338.

Wu, T. F., C. J. Lin, and R. C. Weng. 2004. Probability estimates for multi-class classification by pairwise coupling. *The Journal of Machine Learning Research* 5:975-1005.

Xu, J., R. Tao, Z. Xu, and M. T. Bennett. 2010. China's sloping land conversion program: does expansion equal success? *Land economics* 86(2):219-244.

Zandbergen, P. A., and S. J. Barbeau. 2011. Positional accuracy of assisted GPS data from high-sensitivity GPS-enabled mobile phones. *Journal of Navigation* 64(3):381-399.

Zhang, L. 2012. *In search of paradise: Middle-class living in a Chinese metropolis*. Ithaca: Cornell University Press.

Table 1. The Landsat TM/ETM images used in this study

	1980s		1990s				2000s										2010s			
	88	89	92	93	94	96	00	01	02	03	04	05	06	07	08	09	10	11	12	13
Jan						6		21			6 30		3 19		9		30		20	6 22
Feb							20		9	28		9 25	20	23		12	15		5	7
Mar	30	9 25					23			16	2		8 24	11	29		3 19			11
Apr							24		30	9 25		6	1			17		15		20
May											21		19				6			22
Jun								14								4				
Jul																				
Aug			16																	
Sep							15					13		19						
Oct					25	30			7											
Nov							2	21		19			3 27		8	3 11			19	
Dec				24				23							10	29			21	

Table key:
23: Landsat 7 ETM scan line corrector-off

The images with large cloud coverage (>10%) were omitted. The top row indicates the year of the image that was taken, and the left column indicates the month of the image that was taken. The number in the table indicates the day of the image that was taken.

Table 2. Data description for GIS data other than remotely sensed imagery

	POIs[1,2]	Road network[2]	Social media	Real estate listing record[3]	Real estate transaction record[3]
Sources	Tencent Map	OpenStreetMap, Google Maps, Tencent Map	SINA Weibo	Fang.com	Fang.com
Total number of records	192,468	14,972 km	498,119	1,809	1,393 (in 397 communities)
Attributes in a record	Name, General business type, Specific business type, Extended specific business type	N/A	Post id, user id, sent time, sent device[4], post content, number of user's follower, number of user's following, number of user's total posts, user's gender, user's home province id, user's home city id	Community name, the neighborhood of the community, the district of the community, address, price per square meters, property type, built year, name of developer, building type, total building size, total land size, occupant units, total units, monthly property fee	Community name, transaction date, total transaction price, price per square meters, number of bedroom, number of living room, unit size in square meters, total stories of the building, unit orientation
Example of a record	"Dali Restaurant", Food, Chinese food, Yunnan food	N/A	369597952372XXXX, 204702XXXX, Saturday Apr 5th 2014 12.9pm, Sent from iPhone, "Heavy raindrops fall like tears on Qing Ming. It's all over the windshields of our car", 1369, 219, 1687, female, 13, 10	A beam of sunshine Community, Chuanfang neighborhood, Xishan district, 328 W. Ring Rd, 9753 Yuan/ m^2, Residential community, Jan.1st 2009, Yunnan Xintai real estate, multi-stories building, 84,000 m^2, 78,000 m^2, 247, 261, 1.1 Yuan/ m^2·mon	Wanxin Yinxiang Community, Apr. 10th, 2014, 890,000 Yuan, 9,766 Yuan/m^2., 2 bedrooms, 2 living rooms, 91.13 m^2, 16, south-north orientation

[1] The POIs in our study area in Tencent Maps are superior to those in OpenStreetMap (i.e., 772 in one category) and Google Maps (i.e., 17,590 in one category). To avoid redundancy, we only use Tencent Maps here.
[2] After automatic georeferencing (GCJ02 to WGS84), by our rough estimation, the road network and POI in our study area on Google Maps horizontally offset approximately 131 meters eastward, and 335 meters southward. Those on Tencent Maps horizontally offset 144 meters eastward, and 341 meters southward. We neutralized these offset.
[3] I geocode real estate listing and transaction record with POIs. I address them as online house trading record in the paper.
[4] Unlike Twitter, SINA Weibo makes differences for a variety of mobile devices. I code the price for the most frequently used 145 devices in my study period

Table 3. Major types of land covers and urban land use categories

Land cover types	Major land use categories	Land use sub-categories	Detailed land uses
Built-up areas[1]			
	Manufacturing		Factories
	Transportation		Airport, railway station, toll stations, river dam, road appurtenances, etc.
	Office		Corporations, companies, banks, government buildings, office buildings in schools and hospitals, etc.
	Commercial		Shopping malls and shops, restaurants, night markets, KTVs, car shops, fitness clubs, salons and beauty services, etc.
	Recreational		Convention center, museum, library, buildings in park and zoo, etc.
	Other		Farm house, storage, unused urban land, etc.
	Residential		
		Gated communities	High-rise apartment building complexes, villas, or a large scale of apartment buildings that are often gated and guarded, with few exclusive living services, such as hair salon, spa, day care, etc. suited
		Ordinary communities	6-8 stories Soviet Union style apartment buildings that was once built by work units now is accessible to the public, with the lower levels as stores and restaurants
		Urban slums	1-3 stories small buildings with street level as shops and eat stall, or storages, temporary houses, etc.
Natural elements[2]			
Forest	Urban trees		Street trees, urban park forest, etc.
Grassland	Urban green space		Lawn, garden, urban farm, etc.
Water	Water		Lake, manmade pond, open sewage, etc.
Exposed Soil	Undefined or unused land		Construction site, waste landfill, mine pit, unused land, etc.

[1] Built-up areas here refer to the land surface that is manufactured and can be detected using remote sensing images, regardless of specific human use. In contrast, major land use categories refer to the human use of urban land, roughly corresponding to the major categories of zoning ordinance and urban planning practices. I propose to further classify one major category—residential use into three sub-categories.

[2] Natural elements co-exist as various detailed land uses in urban environment. However, I do not discuss them in detail in this research.

Table 4. Confusion matrix of urban land use major categories accuracy assessment

	Reference								
	Manufacturing	Transportation	Office	Commercial	Recreational	Other	Residential	Total	User's accuracy
Classified									
Manufacturing	15	3	10	6	5	6	8	53	28.3%
Transportation	1	27	3	3	1	3	3	41	65.9%
Office	1	2	31	4	0	4	4	46	67.4%
Commercial	1	1	2	58	1	0	4	67	86.6%
Recreational	1	1	3	0	9	0	4	18	50%
Other	2	1	3	7	2	36	1	52	69.2%
Residential	5	10	9	11	8	3	235	281	83.6%
Total	26	45	61	89	26	52	259	558	
Producer's accuracy	57.7%	60.0%	50.8%	65.2%	34.6%	69.2%	90.7%		
Overall accuracy		73.6%		Kappa coefficient		0.6329			

Figure 1. Study Area

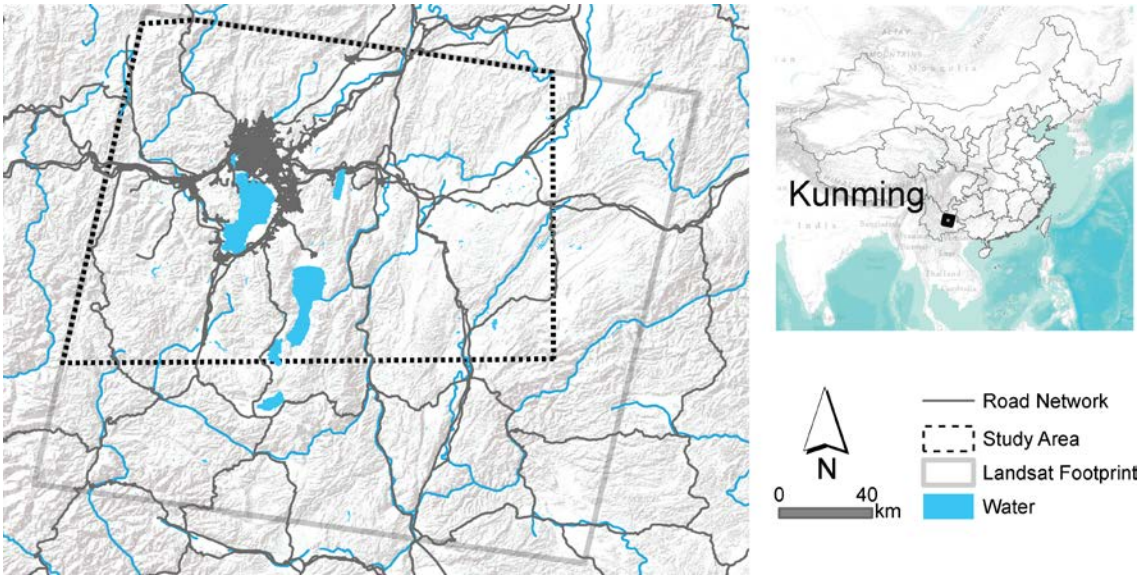


Figure 2. Workflow of land use classification

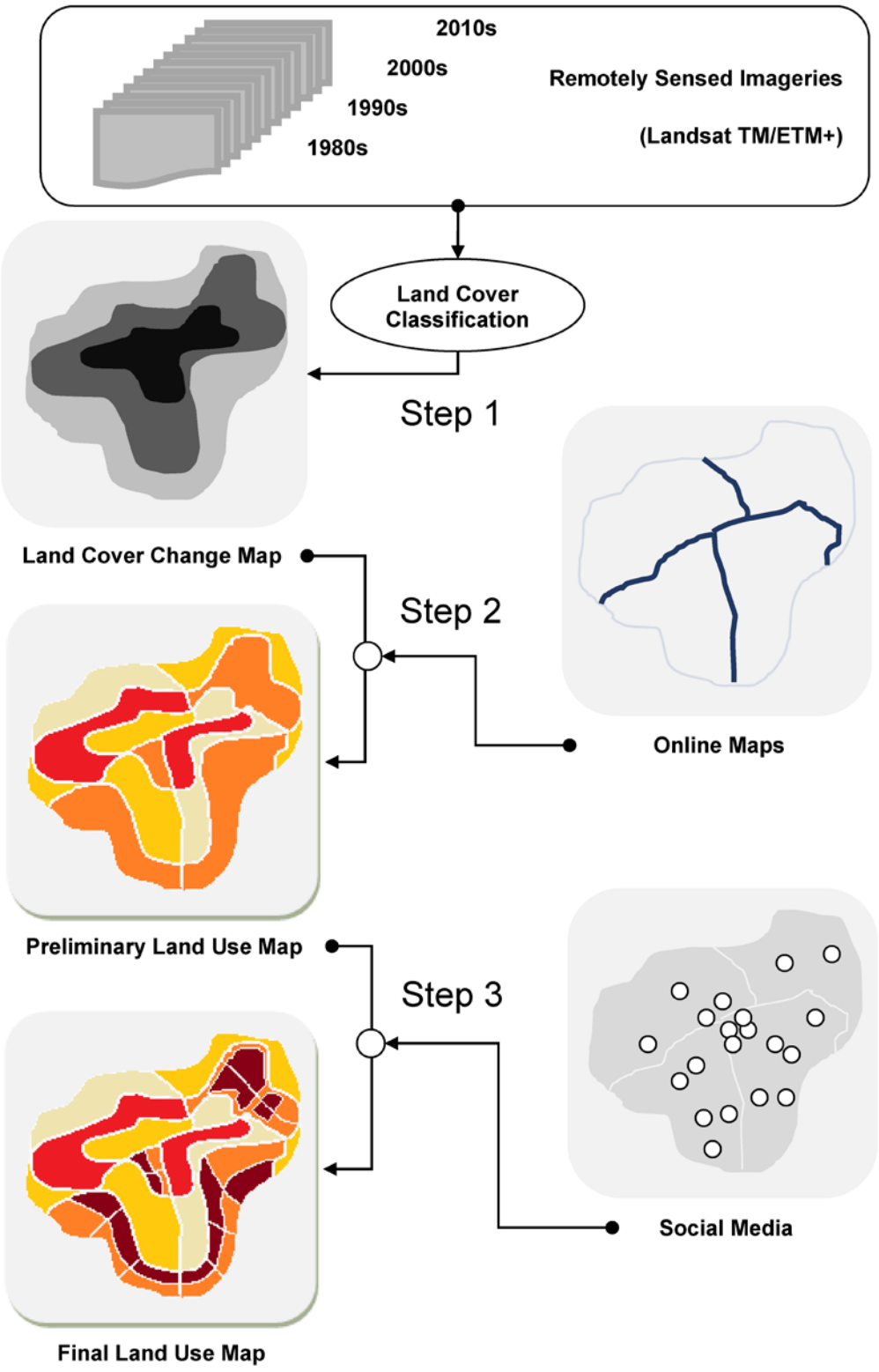


Figure 3. Layout and views of typical land uses

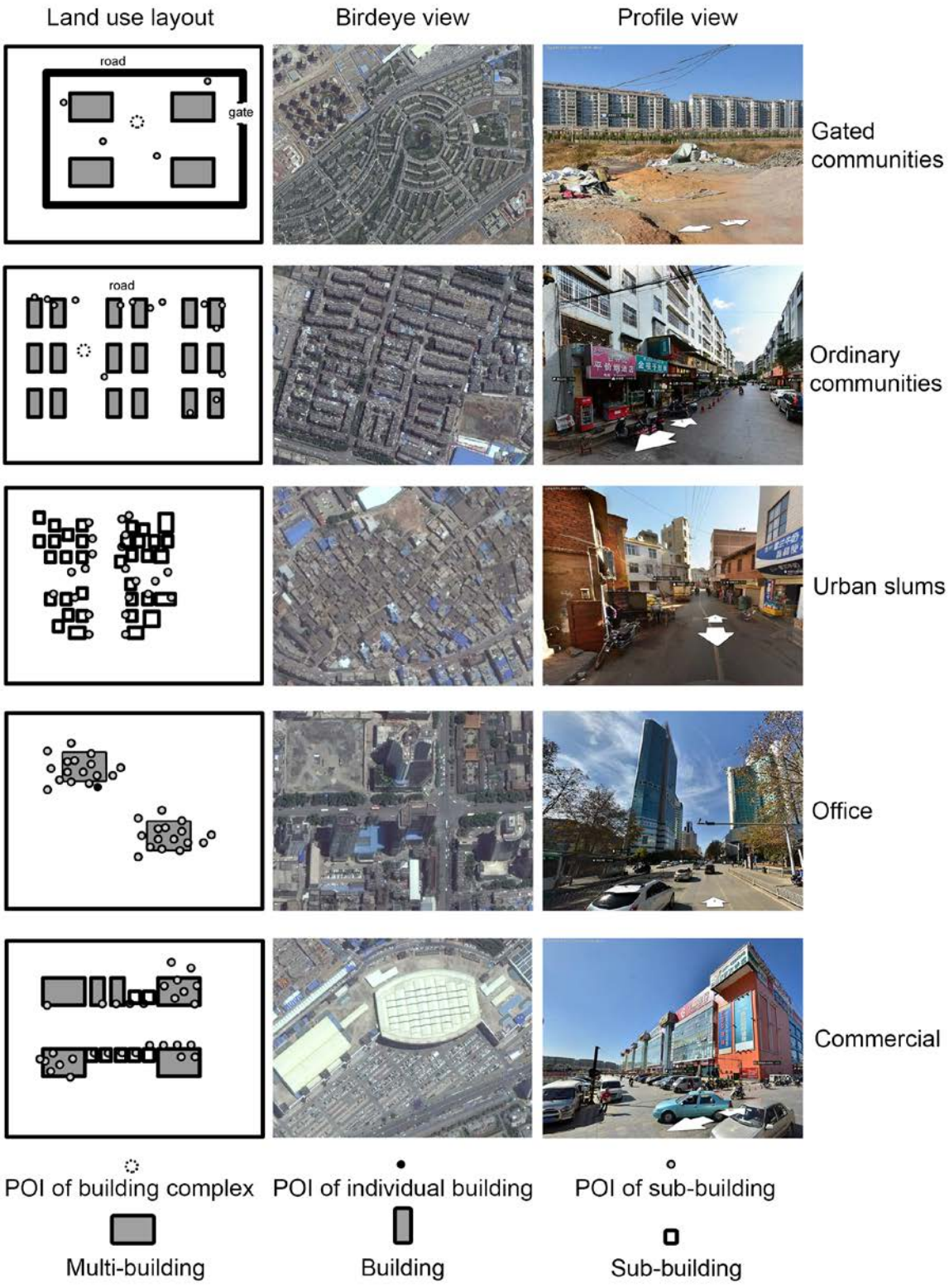


Figure 4. Land cover change in three sample areas

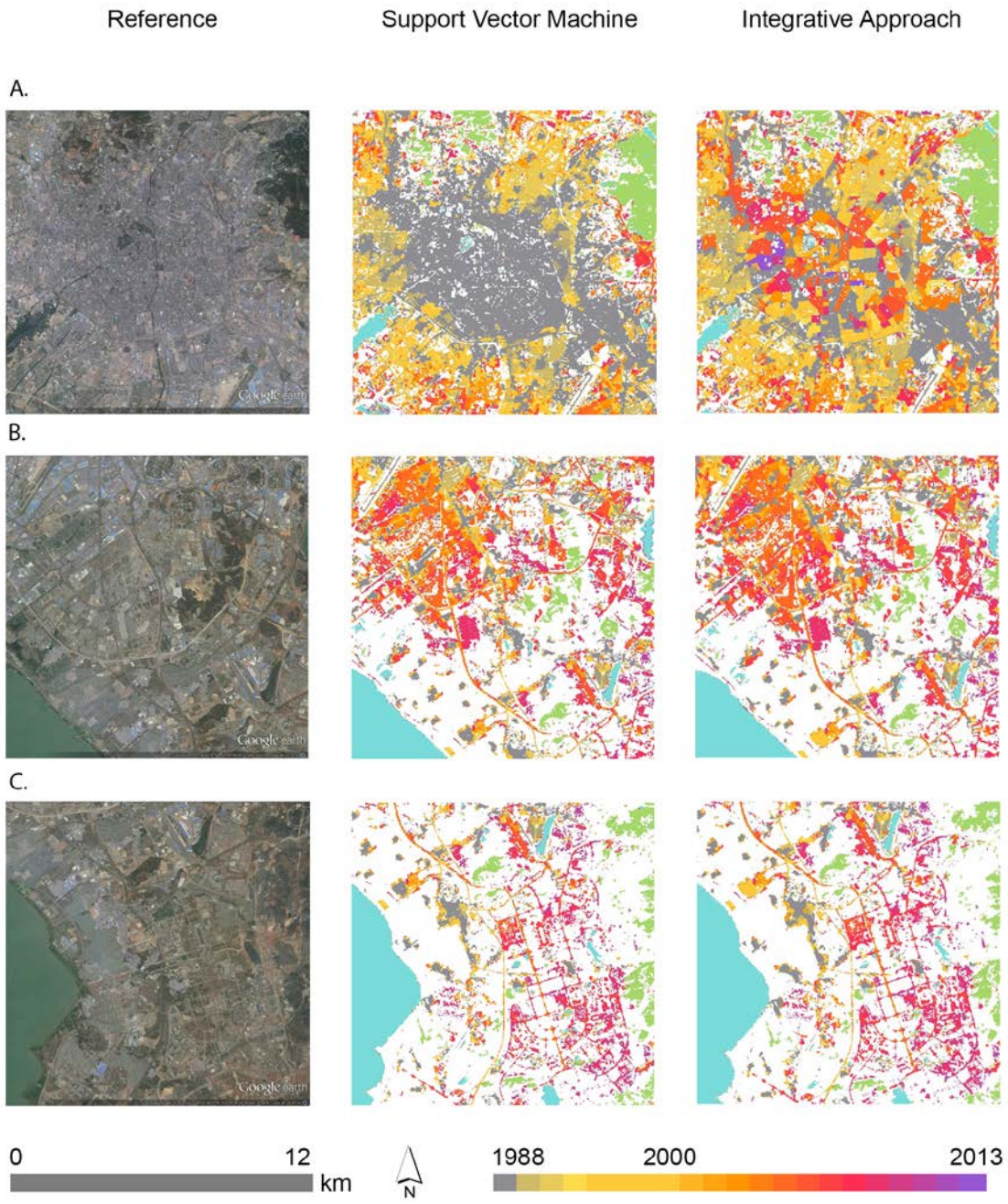


Figure 5. Maps of urban land use in major categories and sub-categories

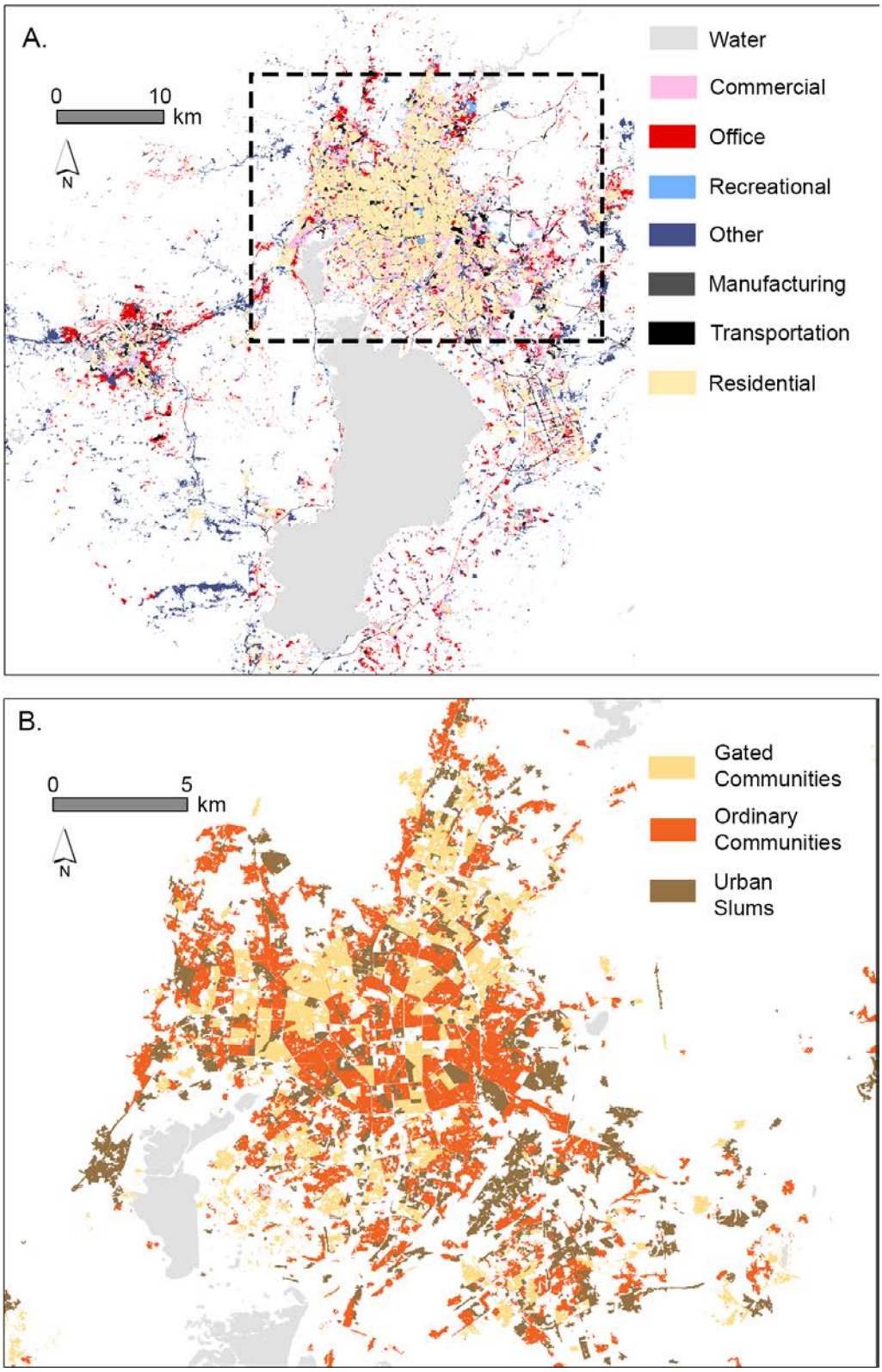
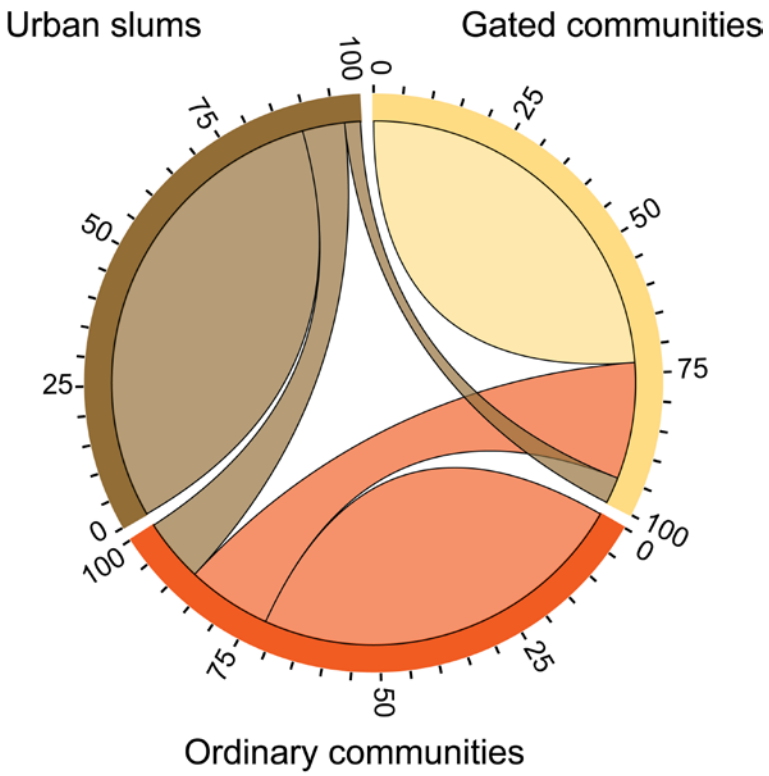


Figure 6. Chord diagram of classification accuracy in residential sub-categories



PUTTING CENSORSHIP IN ITS PLACE: GAUGING BOUNDARIES OF AUTHORITARIAN CONFIDENCE BY ANALYZING GEOTAGGED POSTS IN CHINESE SOCIAL MEDIA

1. Introduction

As recently as a decade ago, the technological revolutions that made real-time exchange of information and opinions among ordinary citizens around the world newly possible and widely accessible seemed to pose a serious challenge to authoritarianism. Yet, even as "pundits, policymakers, and the popular press" (Morozov 2011, 329) noisily proclaimed a social media triumph over clumsy dictators, the authoritarian response was evolving from blunt suppression to a smarter mix of censorship, surveillance, and propaganda. Political science, while never a main force of cyber-utopianism, now better grasps how authoritarian rulers can and do exploit new information and communication technologies to promote a fundamental interest in political stability (Zheng 2008; Greitens 2013; Noesselt 2014; Gunitsky 2015). Moreover, our scholarly understanding of how "networked authoritarianism" (MacKinnon 2011) works in different authoritarian contexts has grown, to take account of the new information management.

Nowhere is the authoritarian machinery to discover, shape, and censor online public talk more developed than in China, under a highly adaptive communist single-party regime. The world's most advanced national firewall selectively blocks access to foreign content (OpenNet Initiative 14 April 2005). A cyberpolice of perhaps 30,000 (Chen and Ang 2011) carries out fulltime surveillance. Hundreds of thousands of paid propagandists (Chen et al. 2011; Henochowicz 2014; Han 2015) flood the cyberspace with political messages that redirect rumors and amplify regime support. These human armies complement the operation of dozens of software products that filter and delete "harmful" information, viewpoints, images, and language (Bamman, O'Connor, and Smith 2012; Fu, Chan, and Chau 2013; Ng 2013; Zhu et al. 2013).

With this machinery, China's rulers manage a huge online population: 668 million netizens in mid-2015, with 89 percent of them accessing the internet on smartphones (China Internet Network Information Center 2015).

We analyze temporal and spatial clustering of some 6.7 million geotagged posts on Sina Weibo, China's sophisticated version of Twitter, to answer the following question: How do powerful authoritarian rulers, obsessed with stability, manage netizens who engage in online political talk (including critical talk) and happen to cluster in a common physical space? In particular, how sensitive are rulers to potentially mobilizing focal points?

The Chinese state is experienced and capable at managing actual outbursts of unsanctioned grassroots collective action—although it evidently prefers to prevent them.[3] Our point of departure is the novel research of King, Pan, and Roberts (2013, 2014) that finds Chinese censorship of the social media aims not to prevent criticism but to prevent collective action, by deleting netizen talk about its actual occurrence, past or future. Put differently: Censorship of Chinese cyberspace reacts to netizen attention to where other netizens (and non-netizens) organize themselves politically to locate in a common physical space. This finding is consistent with a view in the literature that autocrats most fear a coalition of "like-minded people" gathering together in some place to challenge "the regime's stability mantra" (Sullivan 2014). King, Pan, and Roberts arrive at their conclusion through analysis of social media content—but we see its clear implications for the role of *place* in social media censorship. Conceptually and methodologically, we walk through the door they open. We leverage visual

[3] A growing empirical literature documents the Chinese effort to promote social stability by improving local government responsiveness. See, for example, Distelhorst and Hou 2014; Manion 2015; Chen, Pan, and Xu forthcoming; Meng, Pan, and Yang forthcoming.

and statistical discernibility of spatial clustering of posts (including critical posts) about politics, posted in Beijing in 2014−2015, to measure Chinese official sensitivity to unorganized clustering of like-minded ordinary netizens, neither activists nor celebrities, in physical space. We use this measure to go a step further in gauging boundaries of authoritarian confidence about political control of Chinese society. In addition to our dataset of Sina Weibo posts, we draw on original survey data from a probability sample of 2,610 Beijing residents, conducted contemporaneously with our collection of posts.[4]

Even after several catalytic events, we find visually and statistically discernible clustering of netizen political posts, including critical posts, but no criticism posted at Tiananmen Square. We conclude that Chinese rulers are confident enough to permit unorganized spatial clustering of political talk on Sina Weibo, the most potentially subversive form of Chinese social media. This confidence extends to most politically sensitive times and politically sensitive places. However, some places (Tiananmen Square, for example) are focal points that resonate politically with rulers and netizens alike. Especially at politically sensitive times, these places themselves function as (censurable) media.

This article also has broader conceptual and methodological implications. Conceptually, the study of social media has taken the end of geography as premise. Social media connect

[4] The survey is the spring 2015 wave of the Beijing Area Study (BAS) longitudinal survey, conducted annually since 1995 by the Research Center for Contemporary China at Peking University. We designed, for our use here, survey questions on internet use. We use the survey data to assess potential challenges to the external validity of our findings. For a description of the BAS project, see Shen, Yang, and Manion 2010.

hundreds of millions of strangers to one another in cyberspace, which seems to make vast physical spaces irrelevant. Methodologically, in identifying clusters of messages to study, volume bursts direct scholarly attention to points in time as keywords direct attention to communities distinguished by content. In this article, we bring geography to the study of social media. We argue that places themselves can function as media, through which netizens (and rulers) communicate messages. This is especially important where overt political coordination faces big obstacles, as in China. It explains why some places are censurable—and, as we show here, censured by authoritarian rulers.

The remainder of the article is organized as follows. Section 2 introduces Sina Weibo. Section 3 describes our data and explains our use of geotagged posts. Section 4 summarizes how we identify and classify posts for study. Section 5 is the core of the article. It presents, in tests of three sets of hypotheses, our findings from progressively stricter versions of a theoretical claim about an observed sense of regime vulnerability: namely, that netizens clustered in physical space and communicating about politics are a threat to social stability. Section 6 concludes.

2. Sina Weibo

We analyze posts on Sina Weibo, China's version of Twitter. Twitter itself has been blocked in China since July 2009. Sina, China's largest commercial news and entertainment portal, launched its microblogging service in August 2009. Sina Weibo now so dominates the Chinese microblog space that *weibo* (微博, Chinese for microblog) effectively means Sina Weibo. The platform has many features that Twitter lacks: semi-threaded comments, polls, games, photo and video posting, instant messaging, and community portals. Moreover, because a single Chinese character can constitute a complete word, the brevity requirement for posts is not

so constraining.[5] As artist activist Ai Weiwei (2011, 241) puts it: "In Chinese, 140 characters is a novella."

With its many features and without competition from Twitter, Sina Weibo is "the undisputed first source for real-time information in China" (Ng 2013, xiv). Its sophisticated technology and market dominance make it an ideal vehicle for netizens to exchange critical political views. For the regime, the challenge it poses is potentially far more destabilizing than that posed by traditional social media like bulletin board systems or blogs. Unlike these media and the increasingly popular messaging system WeChat, access on Sina Weibo is not restricted to a defined audience. Anything expressed on the platform can be easily, rapidly, broadly—in short, virally—transmitted through reposting. A Chinese Academy of Social Sciences study (Xinhua 24 July 2015) finds Sina Weibo the source of most "rumors": 59 percent, compared to 2 percent for traditional social media.

The purpose of much Chinese online communication (as in other countries) is entertainment (Leibold 2011). Even so, as a demographic, netizens command regime attention. Compared to Chinese not online, netizens are more highly educated, politically opinionated, critical of the regime, and inclined to collective action (Lei 2011). For authoritarian rulers who prioritize social stability above all, Sina Weibo's potential to coordinate officially unsanctioned collective action is worrisome. It is *ipso facto* worrisome to bosses at Sina too, as the media giant bears ultimate responsibility for the content and consequences of whatever appears on its microblog space. Sina's head editor describes it as "a big headache" (Feng 13 June 2010). Sina

[5] Chinese characters usually express concepts in one or two characters. By the estimates of linguists that Hannas (1997) surveys on this issue, single-character words account for 30 or 40 percent of contemporary Chinese writings, and words of three characters or more account for 2 or 3 percent. This makes a Sina Weibo post of 140 characters roughly equivalent to a tweet of 81−86 words.

must integrate concerns of China's rulers into its business routines at the same time as it pursues more user accounts to bring in more revenues from advertising. Sina employs its own labor force of censors (Epstein 14 May 2011) and operates its own fine-grained automated filtering software to rapidly delete posts with politically sensitive content (Bamman, O'Connor, and Smith 2012; Fu, Chan, and Chau 2013; Ng 2013; Zhu et al. 2013). Sina works to strike a balance, conducting "just enough censorship" to satisfy legal and political mandates, without being so intrusive as to drive users away: "If truly innocuous posts disappeared with any regularity, Weibo's users might defect to a competing service" (Zhu et al. 2013). For different reasons, China's rulers must also worry about false positives. Excessive censorship triggers cleverer self-censorship (Chen et al. 2013). In an information-poor institutional environment, with significant incentives for ordinary citizens to falsify preferences and for local agents for falsify achievements, social media like Sina Weibo provide valuable intelligence on public sentiments and the performance of officials far from Beijing; indeed, in China, online venting sometimes yields personnel or policy changes (Noesselt 2014; Hassid 2015).

The posts in our dataset have already undergone whatever censorship Sina exercises. We know Sina's censorship discriminates across users. For example, it targets users who post frequently about politics (Fu, Chan, and Chao 2013; Zhu et al. 2013). It also targets geographically. For example, it deletes more than half the posts originating in Tibet and Qinghai that contain known politically sensitive terms; the deletion rate for comparable posts from Beijing netizens is 12 percent, lowest of all regions in China (Bamman, O'Connor, and Smith 2012).[6] University of Hong Kong Weiboscope data reveal a censorship rate below 1 percent for

[6] For a regularly updated list of these terms (over 9,000 now), see China Digital Times,

the project's very large collection of posts—political, politically sensitive, and otherwise—posted by some 350,000 popular microbloggers.[7] Our 2015 BAS survey findings suggest the rate is lower than this for the Beijing netizens we study here.[8]

3. Data

The Sina Weibo Nearby Application Programing Interface (API) permits us, using a programming interface with customized parameters, to access a database that stores all geotagged posts. Geotagged posts are those posted on a device with software enabled to identify the location of the device at the time of posting. The software is installed on mobile devices such as smartphones, but not on laptop or desktop computers, for example.[9] Netizens can disable geotagging for any particular post or all posts. It is useful to point out that disabling geotagging gives them privacy from other netizens, but certainly not from Sina or (by extension) Chinese state authorities. We designed and systematically navigated a very fine net of locational points to retrieve all geotagged Sina Weibo posts of Beijing netizens for a 350-day period from July 1, 2014 through June 15, 2015.[10] The resulting dataset amounts to some 6.7 million unique posts.

We retrieve posts from Beijing because netizens in the political capital are highly politically attentive.[11] Beijing is also the second highest-density population of microbloggers in

"Sensitive Sina Weibo Search Terms," http://chinadigitaltimes.net/2013/06/grass-mud-horse-list/.

[7] We compute this from Weiboscope data at http://weiboscope.jmsc.hku.hk/datazip/. For geotagged Weiboscope posts, we compute a censorship rate an order of magnitude lower.
[8] With an average 200 posts per day for the netizens followed in the Weiboscope project, a 1 percent probability of censorship for each post implies an 87 percent chance of encountering censorship on any given day. Among surveyed Beijing netizens who access Sina Weibo, 67 percent surveyed say they have never encountered censorship—suggesting a censorship rate much lower than 1 percent.
[9] Also, some applications (Instagram, for example) do not support geotagging.
[10] See Appendix A on data collection.
[11] For example, in our examination of Sina News online stories from late 2012 to mid-2015, we find Beijing netizens comment more than do netizens from any other region.

China (Fu and Chao 2013) and, as noted above, the least censored. We retrieve geotagged posts because we want to introduce geographical place explicitly to the study of the social media. The location we obtain from posts is from GPS on smartphones, thus highly precise. The positional error is 2 to 25 meters (Zandbergen and Barbeau 2011), enabling us to pinpoint accurately the distribution in Beijing of political talk in cyberspace. A focus on geotagged posts has other advantages too. First, the revealed locations are authentic (not self-reported) and confirm that the posts are in fact sent in Beijing. Chinese netizens posting from overseas often self-report a geographic origin within China; including these posts can seriously bias analysis of political talk in the Chinese social media.[12] Second, China's so-called "50-cent party" of paid online commentators are unlikely to be the source of geotagged posts. Mobile devices are a clumsy means to post batches of propaganda; also, high-volume repeated posts from the same location effectively identify paid commentators as such, defeating their purpose. Also, paid online commentators tend to focus more on traditional social media; this may account for the failure of Yang, Yang, and Wilson (2015) to find evidence of their operation on Sina Weibo in 2012−2013.[13] Third, because of mobile device restrictions on reposting, geotagged posts are original, not reposts. In their extraordinary representative sample of Sina Weibo users, Fu and

[12] We have no precise estimate, but we note that among users in the University of Hong Kong Weiboscope 2012 database who geotag as overseas, only 29 percent self-report as overseas. This excludes self-reported Hong Kong netizens.

[13] We are confident that posts by paid commentators do not figure or figure hardly at all in our dataset. Searching through some 4.8 million netizen comments on Sina News online stories from late 2012 through mid-2015, we find 408 Sina Weibo users whose comments on Sina News suggest they are paid commentators. For example, they adopt distinctive nicknames, never make comments that trigger Sina's censors, and make more than 100 positive comments each on stories that focus on sensitive issues of high politics. We find no posts from any of these users in our Sina Weibo database.

Chao (2013) find reposted messages account for 61 percent of posts. Reposts are often items from celebrity microbloggers or from online news sites (including Sina News), which are subject to more stringent information management than is Sina Weibo. Fourth, we are interested in ordinary netizens. The clumsiness of posting on mobile devices discourages not only paid propagandists but also high-volume, high-follower celebrities. By contrast, netizens with lower incomes tend disproportionately to access the internet from smartphones (Sullivan 2012). Finally, posts from mobile devices are from authentic users, not bots that now flood Chinese cyberspace with automated advertising messages.

Geotagged posts present two possible threats to validity for our investigation. One is the device or application from which a post originates; a second, if the device or application indeed supports geotagging, is whether or not geotagging is enabled. Of course, as noted above, disabling geotagging does not provide netizen locational privacy from Sina or the Chinese state. Even so, we make use of the BAS 2015 survey data to empirically assess whether geotagged post content is less political or its political content is less critical. We analyze political differences between netizens who disable the geotagging function when posting on Sina Weibo and netizens who never do so. Nothing we find suggests our focus on geotagged posts poses a problem for our research.[14]

4. Identifying Posts for Analysis

We focus on political posts and critical political posts, identified in several steps. We began with semi-automatic content analysis, sorting our 6.7 million posts by applying a

[14] We report on this test in Appendix B.

collection of 539 Chinese "political" character combinations, in a highly inclusive sense of the term.[15] Posts that use none of these combinations are highly unlikely to have political content. This step identified about 1 percent as political posts. In a second step, our research assistants visually examined the content of all these posts and manually classified them as political, not political, or unclear.[16] This step removed about 60 percent of posts from the political category, identifying 31,163 posts (0.46 percent of all posts) from 21,239 Sina Weibo users as political.[17] This includes posts about local as well as national politics, policies, officials, and institutions. Figure 1 shows the distribution of political posts, as a percentage of daily posts, across the 350 days studied here.

[Figure 1 about here]

We observe a half-dozen volume bursts of political posts in Figure 1. The highest burst, on December 13, marks an officially sponsored celebration of anti-Japanese nationalism: the inaugural Nanjing Massacre Memorial Day to commemorate Chinese killed by the Japanese in Nanjing in 1937. The second highest, on March 1, is the day after Chai Jing's video *Under the Dome*, vividly documenting China's catastrophic air pollution, went viral—generating more than 200 million views before the authorities shut it down a week later. We analyze these two volume bursts in the next section.

[15] We developed these terms by brainstorming with our mainland Chinese research assistants, who (like us) are attentive to Chinese politics, and from examining news stories in official and commercial Chinese media. Terms include names of major political institutions and political leaders, as well as political expressions currently used online to evade keyword filtering.

[16] Our test using about 10 percent of posts to verify coding rule reliability found 80 percent intercoder agreement across two coders, with a Cohen's kappa of 0.61.

[17] Among our 21,239 users with political posts, 33 percent have fewer than 100 followers, 50 percent have 100−1,000 followers, 13 percent have 1,000−10,000 followers, and 4 percent have more than 10,000 followers. For perspective, compare to celebrity microblogger Ren Zhiqiang, who commanded over 37 million followers when the authorities shut down his account in March 2016.

The third highest volume burst occurs on January 1, catalyzed by a stampede of New Year's Eve revelers on an overcrowded Shanghai riverfront, leaving dozens dead. This burst is effectively indistinguishable from the one at the end of June; the June posts mostly focus on China's air pollution, but we know of no obvious catalyzing event. The fifth and sixth highest volume bursts are (like December 13) official political celebrations. National Day on October 1 commemorates the founding of the People's Republic of China in 1949. Constitution Day on December 4 is an inaugural holiday celebrating China's Constitution.

Two smaller bursts, occurring on July 29 and late December, are effectively indistinguishable from one another in volume. We analyze the July volume burst in the next section. It marks the highly publicized announcement of a communist party investigation into the corruption of Zhou Yongkang. Before his retirement in 2012, Zhou was at the apex of power as China's domestic security chief and one of nine Communist Party Politburo Standing Committee members.

5. Analysis

We treat visual and statistical discernibility of spatial clustering of political posts as an indication that the like-minded netizens of interest to us do cluster spatially, but this point is not in itself of particular analytic interest. Rather, if we can observe that political posts in our dataset of already censored posts discernibly cluster, it interests us because of implications about (absence of) censorship of netizens who coincide in physical space and gather in cyberspace to talk about politics. In particular, we look in our analysis for spatial clustering of political posts, including critical political posts on politically sensitive dates.

Given the official obsession with social stability, the sensible theoretical place to start is a claim of observed regime *vulnerability*. That is, China's rulers sense a regime threat in the

unorganized clustering in physical space of netizens talking about politics; blunt censorship of Sina Weibo political posts reflects this obsession. We formulate our null hypothesis accordingly, to reflect censorship: indiscernibility of spatially clustered political posts. Conversely, to the extent that we do discern spatial clustering of the political posts in our dataset of already censored posts, we reject the null in favor of an alternative hypothesis, a *confident* regime.

We analyze different situations and catalytic political events to test progressively stricter versions of the null. Variation in discernible spatial clustering of political posts across situations and political catalysts is our way of mapping observed vulnerability of Chinese rulers vis-à-vis the society they govern. It allows us to make inferences about regime sensitivity to focal points that (in principle) have greater than usual potential to mobilize public opinion, perhaps even collective action. We are especially interested in events that seem neither obviously threatening nor obviously unthreatening. For example, we analyze sensitivity to netizen spatial clustering after release of Chai Jing's *Under the Dome*. Places themselves can also be focal points. For example, we analyze sensitivity to netizen clustering at Tiananmen Square.

5.1. Baseline Spatial Clustering

We begin in this section with our 31,163 political posts, by which we identify like-minded netizens of interest, aggregated across our entire period of study. We presume that like-minded netizens do in fact cluster in physical space, as the simple result of socialization or recruitment to workplaces and residential communities that separate Chinese of different socioeconomic statuses, for example. Indiscernibility of spatial clustering of political posts in our entire dataset of posts should (we think) be a relatively easy null to falsify across 350 days: China's rulers ought to be at least confident enough to refrain from censoring these posts, many of which are separated from one another by months. We test this hypothesis by examining the

discernibility of spatial clustering of all posts our research assistants coded as political for the period from July 1, 2014 through June 15, 2015. The distribution is mapped in Figure 2.[18] The small square at the map's center is Tiananmen Square. We analyze posts (and absence of posts) from the square in Section 5.3.

[Figure 2 about here]

The spatial clustering of posts in Figure 2 is visually discernible. Shades of yellow and orange indicate a density of political posts higher than the 0.46 percent average for the 350-day period. We observe that political posts are most densely clustered within the innermost ring road; they are especially densely clustered at the very center of the map, in the orange shaded space where 1−2 percent of posts are political posts. This space includes Tiananmen Square and government offices such as the Ministry of Public Security and Beijing City Government. That is, Beijing's political center is also the center of political talk in cyberspace. Further away from the city center, the yellow shaded spaces show other (less dense) clusters of political posts. For example, posts spatially cluster just inside the third ring road in the city's northwest suburbs; this cluster is an upscale residential community for those working in the nearby high-technology district and also a popular rental apartment district for university students from nearby campuses, such as Peking University and Renmin University.

The clustering of political posts visible in Figure 2 is statistically discernible too. We draw a random sample of posts to examine the degree of clustering at different spatial scales. We use the Ripley's Cross-K function to estimate clustering at scales from 0 meters to 40 kilometers. The posts are indeed clustered (not statistically random or dispersed) in Beijing's physical

[18] We can easily place a post in a single building that we can identify on a Google Earth map. We aggregate sufficiently in our presentations in this article to protect human subjects.

space.[19]

The visual and statistical discernibility of spatial clustering mapped in Figure 2 suggests that China's rulers are confident enough to allow netizens who find themselves clustered in their residential and work communities to express themselves as though they have a right of virtual assembly for everyday political talk. The lack of support for the null in this situation, with netizen posts aggregated over a long span of time, does not surprise us. We present it here as a baseline before introducing a shorter time frame and a range of political catalysts to test hypotheses of greater substantive interest, which we turn to immediately below.

5.2. Spatial Clustering at Particular Times

In this section, we organize our hypothesis testing by examining the spatial clustering of political posts over four short chunks of time, each intersected by a potentially mobilizing event we identify by our visual examination of volume bursts in Figure 1 or (in one case) knowledge of Chinese politics. The events are reported in chronological order in Table 1.

[Table 1 about here]

Chronologically the last of the four events, June 4, 2015, anchors our expectations about observed regime vulnerability. It marks the anniversary of the 1989 military suppression of the astonishing months-long occupation of Tiananmen Square by hundreds of thousands of protesters. In China, no day is more politically sensitive than this anniversary. We note no volume burst on June 4 in Figure 1. This is because cyberspace references to the event are

[19] At all degrees (i.e., from 0 meters to 40 kilometers), the random sample of political posts for the 350-day period exhibits more (and different) clustering than does a same-sized random sample drawn from all (political and non-political) 6.7 million posts. In this section and others, we map posts (not users) because it is more straightforward, but we also randomly sampled one post from each user who sends more than one post. Spatial clustering of users is also visually and statistically discernible.

swiftly and predictably censored; netizens know this and self-censor by employing sly language morphs or (more commonly now, decades after the event) avoiding the topic entirely.[20] The authorities prepare obsessively before June 4 anniversaries: augmenting security around the square, placing known troublemakers under house arrest or transporting them out of town, and distracting ordinary Chinese with entertainment. Internet censors sometimes ban mundane terms such as "tomorrow" and "today" as June 4 approaches and passes. We expect official high anxiety to be evident in censorship of political posts in the days preceding, on, and soon after June 4.

We are most interested, however, in a more nuanced test of the boundaries of observed regime vulnerability, which is provided by the other three events. None is nearly as politically sensitive as the June 4 anniversary, but all three are provocative focal points with some potential to mobilize public opinion against the regime, perhaps even coordinate collective action.

The announcement on July 29, 2014 of the party's investigation into Zhou Yongkang for corruption clearly reflects politics at the apex of power, which is always politically sensitive. China is embroiled in the longest anticorruption campaign in its communist party history. The campaign is a signal achievement for Xi Jinping, who launched it soon after his ascension to the party leadership in 2012. It has felled more officials and more senior officials than any in party history—and the party massively publicizes its campaign in the official media. Yet, exposing the corruption of an official as senior as Zhou can encourage greater cynicism about the rot in elite politics and about the campaign as window dressing for political purge.

[20] For examples of language morphs, see China Digital Times, "Sensitive Sina Weibo Search Terms," http://chinadigitaltimes.net/2013/06/grass-mud-horse-list/.

The inaugural Nanjing Massacre Memorial Day on December 13, 2014 is an officially sponsored event that, in principle, poses little threat to social stability. A celebration of nationalism, it is intended to unify Chinese by recalling an especially brutal experience of national humiliation. At the same time, commemorating any such historical event is a two-edged sword. Ultra-nationalist blogs are popular in China, and anti-Japanese nationalism has exploded in violent protest in recent years.[21]

Chai Jing's video, *Under the Dome*, was released on February 28, 2015. It pointedly documents China's devastating air pollution; it is apocalyptic in its vision of China's environmental future. Unlike the Nanjing Massacre Memorial Day and the announcement of the party's investigation of Zhou Yongkang, it qualifies as a political shock. It sets out a policy failure of huge proportions. Certainly, ordinary Chinese are keenly sensitive to the environmental degradation around them. Indeed, the issue routinely generates mass protests, in both the countryside and cities. The timing of the release also seems provocative. In early March, thousands of National People's Congress and Chinese People's Political Consultative Conference delegates descend on Beijing for ten days of the annual "two meetings." These days are times of heightened political sensitivity in Beijing, especially at the city center where the delegates meet. Yet, *Under the Dome* was released not only on the Chinese equivalent of YouTube but also on the official website of the communist party daily. Also, China's new Minister for Environmental Protection added to the implicit official endorsement with his own praise for the documentary the day after its release. Chai herself is a former reporter for Chinese Central Television, hardly the credentials of a political activist.

[21] Indeed, King, Pan, and Roberts (2013, 2014) treat nationalist sentiments as analytically equivalent to references to collective action events in their study of censorship.

Do China's rulers seek to engage or suppress political talk in cyberspace of netizens who happen to cluster together in physical space soon after potentially mobilizing political events? We present results of our mapping exercise in two columns of maps in Figure 3.[22] As above, the null hypothesis is observed regime vulnerability, that is: indiscernibility of spatial clustering of political posts, in this case in the 72 hours after each event, mapped in column 2. The map shading follows the same logic as above. Political post density exceeds the 350-day average in spaces shaded yellow and (in this figure) several shades of orange and a deep red.

[Figure 3 about here]

Most striking in Figure 3 are the expanses of white space in the two maps at the top of each column, reflecting extraordinary official censorship and self-censorship in the days before, on, and after June 4. Netizen political talk, to say nothing of spatially clustered political talk, is virtually off the table. The political posts we read from the city center on and after 4 June confirm the extreme sensitivity of what we already know is the most sensitive of political anniversaries, more than a quarter-century after 1989. For example, a political post from June 4 is quite pointed, even as it deliberately avoids use of politically sensitive keywords:

- It is a complicated day for this country. The course of history washed away dust and solidified the moments that in any case should not be forgotten. For those who made efforts for the lives of people, we salute them. There is always a force to support the nation's forward movement, and I believe the most beautiful day will eventually arrive. Then, we will be able to listen to the purest sound of the bells of the republic!

[22] We are investigating political talk at politically sensitive times. Most political posts mapped here directly comment on the event they bracket: 30 percent of political posts after the announcement of the investigation of Zhou Yongkang, 76 percent after the beginning of the Nanjing Massacre Memorial Day, and 75 percent after the release of *Under the Dome*. All critical posts (mapped and analyzed below) are directly relevant comments.

Censorship on June 4 is spatial censorship too, cordoning off Tiananmen Square, as this post indicates:

- Today the security check is very strict. To get onto the square, we must walk around to the end of long lines. Everyone's ID must be documented, and body searches are stricter than at the airport. Bags must be thoroughly searched after the security scan. Notebooks must be checked page by page. Dashuai [nickname of person who evidently accompanies microblogger to square] asked [the security officer]: "Is this what you do every day?" The security [officer] answered: "Today is special." Many people my parents' age are dressed all in black standing in line, some even wearing mourning clothes.[23]

Also striking in Figure 3 is the evidently catalyzing effect of the other three events on netizen political talk. In the 72 hours that follow each, white maps acquire yellow, orange, and red shading. In many spaces, observed netizen political talk exceeds 4 and even 8 percent of all posts. *Under the Dome* is particularly catalyzing, which of course we know already from examination of volume bursts. Tiananmen Square and the city center are hot spots for netizen political posts. More to the point for our investigation, the clustering of political posts is visually and statistically discernible.[24]

We infer from Figure 3 that China's rulers are confident enough to permit unorganized spatial clustering of netizens for political talk, even after the occurrence of political events with some potential to mobilize public opinion, perhaps even to coordinate collective action against official politics, policies, officials, or institutions. This includes anti-Japanese nationalist political

[23] We note ambiguity in the last line of the post; it can be read as a straightforward reference to mourning for Mao Zedong by visitors to Mao's mausoleum, which is on the square.
[24] Posts in the 72 hours after June 4 are too few to examine for statistical evidence of clustering.

talk. At the same time, the contrast with the June 4 anniversary starkly shows this confidence is not unbounded. China's rulers have absolutely no interest in encouraging netizen political talk soon before, on, or after June 4.

We found a sizable number of political posts voicing strong criticism. A few examples illustrate:

- #National [Nanjing Massacre] Memorial Day# The nation-state is really just a tool for domination. What really connects us is race, if we don't want to lose sight of the past then we must perceive it clearly.
- Everything about Zhou [Yongkang]'s past is exposed suddenly, overnight, he is guilty of unpardonable evil. What about party regulations and the law, there is evidence for [violation of] both. How then can such an evil type rise to the Politburo Standing Committee? What a huge flaw in the system, isn't it a disgrace?
- Worship of the head of state, the sovereign, the heavenly emperor, the dictator has reached dangerous levels. [I'm] feeling ill.

We asked our research assistants to visually examine all political posts for the 72 hours following each of the four events, classifying posts as critical or not.[25] The panel of maps in Figure 4 maps the density of critical political posts. Because critical posts are a small subset of political posts, we use bigger categories to map post density, but the logic of map shading is unchanged.

[Figure 4 about here]

[25] Posts classified as not critical are not necessarily supportive; they may be neutral or unclear in orientation. Our coding is fairly conservative. For example, we do not code posts referring to a "blue sky" as *ipso facto* critical, even if posted in the 72 hours after release of *Under the Dome*. Posts with references to "APEC blue" or "two meetings blue" are coded as political, however. See below.

The fairly white map of June 4 in Figure 3 is a totally white map in Figure 4, reflecting the totality of censorship (including self-censorship) surrounding this event. White space dominates the Nanjing Massacre Memorial Day map; shaded areas reflecting spatial clustering of critical posts about this event are few and also far from the city center. By contrast, the map of the 72 hours after announcement of the investigation into the corruption of mega-tiger Zhou Yongkang features over a dozen spatial clusters where critical posts exceed 66 percent of political posts. These posts account for about half of all critical posts for all four events. Moreover, they cluster closer to the city center than do critical posts that follow the other events.

We now know that the authorities shut down *Under the Dome*, deleting it from viewership on the internet seven days after its release. News stories about the documentary were deleted at the same time. This response of China's rulers is not reflected in Figure 4. Indeed, even before deleting the documentary, the authorities began to censor netizen response. Beginning on March 2, netizen comment on *Under the Dome* was prohibited on many streaming websites. This response is not fully reflected in Figure 4 either. In Figure 5 we map the interplay of netizens and authorities at three different 72-hour chunks of time corresponding to this succession of information management regimes in the nine days after the policy shock of the documentary's release. To be sure, the three stages are not as clear cut as we would like. For example, netizen comments were prohibited on some but not all streaming websites as of noon on March 2. Certainly, however, it makes theoretical sense to distinguish three stages, rather than lump the entire period together or ignore changes in information management that began about 72 hours after the documentary's release.

[Figure 5 about here]

Map shading and categories in Figure 5 follow the logic in previous figures. In the first

column, the contrast between the top and bottom maps is most striking. The release of *Under the Dome* is a catalyst for political talk, but the discernible spatial clusters of political talk are fewer and much less dense by days 7, 8, and 9. The second column of maps suggests that the implicit official endorsement of the documentary may have sheltered the regime from political criticism. The big contrast is with the middle map, which suggests a possible counterproductive impact of the onset of censorship. In this second 72-hour period, netizen critical political talk increased. Some examples follow:

- Following APEC blue [the artificially-induced clear skies for the November 2014 Beijing meeting of the Asia-Pacific Economic Cooperation meeting] we now have two meetings blue. 80% of Chai Jing's effort has been wasted. Officials just have meetings and more meetings.
- My friends say that if President Xi would hold more meetings, we could then enjoy more blue sky time!
- Wind scatters the smog, but what if the wind stops? The government claims that reducing smog will not depend on wind! But sunny days only come after wind! The party claims APEC blue will become the new normal. What a "Chinese Dream"![26] Anyway, be happy on such a smoggy day! The more difficult your life, the happier you should live it!

The expanse of white space on the bottom map, replacing dense clusters of critical political posts, most likely reflects the more decisive censorship regime of days 7, 8, and 9.

We note that on all maps in Figure 4 as well as the top and bottom maps in Figure 5, no critical posts appear to originate from Tiananmen Square. The iconic space at the city center

[26] "Chinese Dream" is the signature slogan of top communist party leader Xi Jinping.

appears empty of critical talk in cyberspace. We turn to this immediately below.

5.3. Tiananmen Square

Not all space is politically equal: geography is also a political matter—by design in Beijing especially, with its history as China's political capital for centuries of imperial rule and all six decades of communist rule. Places themselves can function as media, through which rulers (and netizens) communicate messages. Figures 4 above suggests Tiananmen Square, the politically symbolic public square in Beijing's city center, is a megaphone for political talk. It amplifies political messages, so that somewhat politically sensitive posts become highly sensitive posts—simply because they originate with netizens on the square.

Tiananmen Square resonates politically, with rulers and netizens alike. It is bounded on its north by the Forbidden City, home of China's emperors; the imperial gate Zhengyangmen (or Qianmen) is at the square's south. The communist party has claimed the site for its own history. In 1949, Mao Zedong announced the founding of the People's Republic of China from the rostrum of the Gate of Heavenly Peace that demarcates the square from the Forbidden City. In the 1950s, the regime erected the buildings that now define its west and east boundaries: the Great Hall of the People (for congress meetings) and the Museum of Revolutionary History. In 1976, Mao's mausoleum was built, taking up a large area in the square's south. The square's political relevance is accentuated by its proximity to Zhongnanhai, the guarded compound a few blocks west of the Forbidden City, headquarters to the Central Committee and State Council. Major party and government office buildings are close to the square. Because of its proximity to and identification with major sites of Chinese traditional and contemporary political significance, the square also functions as a focal point for ordinary citizens in Beijing. For example, long before 1989, in April 1976, protesters gathered in large numbers at the square to air anti-regime

criticism before their violent suppression by the authorities. In January 1987, the authorities physically blocked attempts by protesters to assemble at the square in support of liberal reforms.

If Tiananmen Square is effectively off limits for political criticism, what are the spatial boundaries of observed regime vulnerability? For example, Beijing's (historically) politically volatile university district is about 20 kilometers northwest of the square. In the eyes of China's rulers, is this a comfortable distance from which netizens may discernibly cluster to engage in critical political talk at times of potentially mobilizing events? To gauge with some precision that "comfortable distance," we estimate logistic models with dichotomous dependent variables that distinguish between netizens who post critical political posts and netizens whose political posts are neutral, supportive, or not clearly critical. The catalytic political events are categorical independent variables. For the Nanjing Massacre Memorial Day and the announcement of the investigation of mega-tiger Zhou, these are simple dichotomous variables; for *Under the Dome*, we create a dichotomous variable for each of the three stages and treat the first 72 hours after the documentary's release as the reference category. As we note above, the timing of three stages is not as clear-cut as we would like, but certainly there are three stages that correspond reasonably well with our three 72-hour periods.

The independent variable of key theoretical interest is distance of the critical post origin from Tiananmen Square. We measure this as "network distance," to take into account actual travel distance from the post point of origin to Tiananmen Square, as though a netizen were taking a taxi on the existing road network to the square.[27] In model 2, we control for netizen influence (number of followers) and netizen involvement in the social media (number of

[27] We estimate network distance using Baidu Direction API. See http://lbsyun.baidu.com/index.php?title=webapi/direction-api.

followings). Results are reported in Table 2; the impact of network distance to the square in model 1 is shown in Figure 6.

[Table 2 and Figure 6 about here]

As reported in Table 2 and shown in Figure 6, the higher the probability that a Sina Weibo user's politically critical post remains discernible in the 72 hours after our catalytic political events, the greater the distance from Tiananmen Square, in both a statistical and visual sense. Network distance to the square, the variable of theoretical interest, is statistically significant. The evidence for the relationship between discernible political criticism and distance from the square is clearly evident in Figure 6. In sum, to paraphrase Marshall McLuhan and in light of the Arab Spring, it appears the *maidan* is the message.

6. Conclusion

Research by King, Pan, and Roberts (2013, 2014) shows that Chinese censorship in cyberspace aims to delete netizen talk about collective action. In this article, we extend the investigation by bringing geography into the study of social media, showing how physical space matters for information management in authoritarian regimes. We find that China's rulers permit clustering of critical talk by ordinary netizens, even if the talk is sometimes subversive, provided it does not occur at sensitive times or sensitive places. On the extremely sensitive anniversary of June 4, for example, we find Beijing netizen are either not permitted or do not dare to voice discontent. In other examples, such as the investigation into the corruption of mega-tiger Zhou Yongkang, the release of Chai Jing's antipollution documentary *Under the Dome*, and the regime-sponsored Nanjing Massacre Memorial Day, we find netizens in Beijing do voice discontent—but, as our analysis shows, their discontent clusters at some distance from Tiananmen Square. What is remarkable about Tiananmen Square at politically sensitive times is

that even without netizen voices of discontent or the presence of a single netizen, it still has the power to be a medium, indeed a megaphone, that China's authoritarian rulers evidently find threatening.

REFERENCES

Ai Weiwei. 2011. *Ai Weiwei's Blog: Writings, Interviews and Digital Rants, 2006–2009*. Cambridge, Mass.: MIT Press. Edited and translated by Lee Ambrozy.

Bamman, David, Brendan O'Connor, and Noah A. Smith. 2012. "Censorship and Deletion Practices in Chinese Social Media." *First Monday*, vol. 17, no. 3. http://firstmonday.org/ojs/index.php/fm/article/view/3943/3169.

Chen, Cheng, Kui Wu, Venkatesh Srinivasan, and Xudong Zhang. 2011. "Battling the Internet Water Army: Detection of Hidden Paid Posters." *arXiv*. doi: 10.1145/2492517.2492637.

Chen, Jidong, Jennifer Pan, and Yiqing Xu. Forthcoming. "Sources of Authoritarian Responsiveness: A Field Experiment." *American Journal of Political Science*. doi: 10.1111/ajps.12207.

Chen, Xiaoyan, and Peng Hwa Ang. 2011. "The Internet Police in China: Regulation, Scope, and Myths." In *Online Society in China: Creating, Celebrating, and Instrumentalising the Online Carnival*, edited by David Herold and Peter Marolt, 40–52. New York: Routledge.

China Internet Network Information Center. 2015. "36th Statistical Report on Internet Development in China." http://www1.cnnic.cn/IDR/ReportDownloads/201601/P020160106496544403584.pdf

Distelhorst, Gregory, and Yue Hou. 2014. "Ingroup Bias in Official Behavior: A National Field Experiment in China." *Quarterly Journal of Political Science*, vol. 9, no. 2: 203−230.

Epstein, Gady. 14 March 2011. "Sina Weibo." *Forbes Asia*. http://www.forbes.com/global/2011/0314/features-charles-chao-twitter-fanfou-china-sina-weibo.html.

Feng, David. 13 June 2010. "Sina Head Editor: Censoring Tweets—What a Pain!" *techblog86*. http://www.techblog86.com/2010/06/13/sina-head-editor-censoring-tweets-what-a-pain/.

Fu, King-wa, Chung-hong Chan, and Michael Chau. 2013. "Assessing Censorship on Microblogs in China: Discriminatory Keyword Analysis and the Real-Name Registration Policy." *IEEE Internet Computing*, vol. 17, no. 3: 42–50.

Fu, King-wa, and Michael Chau. 2013. "Reality Check for the Chinese Microblog Space: A Random Sampling Approach." *PLoS One*, vol. 8, no. 3. http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0058356.

Greitens, Sheena Chestnut. 2013. "Authoritarianism Online: What Can We Learn from Internet Data in Nondemocracies." *PS: Political Science and Politics*, vol. 46, no. 2: 262–270.

Gunitsky, Seva. 2015. “Corrupting the Cyber-Commons: Social Media as a Tool of Autocratic Stability.” *Perspectives on Politics*, vol. 13, no. 1: 42–54.

Han, Rongbin. 2015. “Manufacturing Consent in Cyberspace: China’s ‘Fifty-Cent’ Army.”

Hannas, Wm. C. 1997. *Asia’s Orthographic Dilemma*. Honolulu: University of Hawai’i Press. *Journal of Current Chinese Affairs*, vol. 44, no. 2: 105–134.

Hassid, Jonathan. 2015. “China’s Responsiveness to Internet Opinion: A Double-Edged Sword.” *Journal of Current Chinese Affairs*, vol. 44, no. 2: 39−68.

Henochowicz, Anne. 3 December 2014. “Thousands of Local Internet Propaganda Emails Leaked.” *China Digital Times*. http://chinadigitaltimes.net/2014/12/thousands-local-internet-propaganda-emails-leaked/.

King, Gary, Jennifer Pan, and Margaret E. Roberts. 2013. “How Censorship in China Allows Government Criticism but Silences Collective Expression.” *American Political Science Review*, vol. 107, no. 2: 326–343.

King, Gary, Jennifer Pan, and Margaret E. Roberts. 2014. “Reverse-Engineering Censorship in China: Randomized Experimentation and Participant Observation.” *Science*, vol. 345, no. 6199: 1–10.

Lei, Ya-Wen. 2011. “The Political Consequences of the Rise of the Internet: Political Beliefs and Practices of Chinese Netizens.” *Political Communication*, vol. 28, no. 3: 291–322.

Leibold, James. 2011. “Blogging Alone: China, the Internet, and the Democratic Illusion?” *Journal of Asian Studies*, vol. 70, no. 4: 1023–1041.

MacKinnon, Rebecca. 2011. “China’s ‘Networked Authoritarianism’.” *Journal of Democracy*, vol. 22, no. 2: 32–46.

Manion, Melanie. 2015. *Information for Autocrats: Representation in Chinese Local Congresses*. Cambridge: Cambridge University Press.

Meng, Tianguang, Jennifer Pan, and Ping Yang. Forthcoming. “Conditional Receptivity to Citizen Participation: Evidence from a Survey Experiment in China.” *Comparative Political Studies*. doi: 10.1177/0010414014556212.

Morozov, Evgeny. 2011. *The Net Delusion: The Dark Side of Internet Freedom*. New York: Public Affairs.

Ng, Jason Q. 2013. *Blocked on Weibo: What Gets Suppressed on China’s Version of Twitter (and Why)*. New York: New Press.

Noesselt, Nele. 2014. “Microblogs and the Adaptation of the Chinese Party-State’s Governance Strategy.” *Governance*, vol. 27, no. 3: 449–468.

OpenNet Initiative. 14 April 2005. “Internet Filtering in China 2004–2005: A Country Study.” https://opennet.net/sites/opennet.net/files/ONI_China_Country_Study.pdf.

Shen, Mingming, Ming Yang, and Melanie Manion. 2010. “Measuring Change and Stability over a Decade in the Beijing Area Study.” In *Chinese Politics: New Sources, Methods, and Field Strategies*, edited by Allen Carlson, Mary Gallagher, Kenneth Lieberthal, and Melanie Manion, 236–245. New York: Cambridge University Press, 2010.

Sullivan, Jonathan. 2012. “A Tale of Two Microblogs in China.” *Media, Culture, and Society*, vol. 34, no. 6: 773–783.

Sullivan, Jonathan. 2014. “China’s Weibo: Is Faster Different?” *New Media and Society*, vol. 16, no. 1: 24–37.

Xinhua. 24 July 2015. “What Is the Source of Rumors in the New Media?” In Chinese.

Yang, Xiaofeng, Qian Yang, and Christo Wilson. April 2015. “Penny for Your Thoughts: Searching for the 50 Cent Party on Sina Weibo.” International AAAI Conference on Web and Social Media. http://www.aaai.org/ocs/index.php/ICWSM/ICWSM15/paper/view/10524.

Zandbergen, Paul A., and Sean J. Barbeau. 2011. “Positional Accuracy of Assisted GPS Data from High-Sensitivity GPS-enabled Mobile Phones.” *Journal of Navigation*, vol. 64, no. 3: 381–399.

Zheng, Yongnian. 2008. *Technological Empowerment: The Internet, State, and Society in China*. Palo Alto, Cal.: Stanford University Press.

Zhu, Tao, David Phipps, Adam Pridgen, Jedidiah R. Crandall, and Dan S. Wallach. 2013. “The Velocity of Censorship: High-Fidelity Detection of Microblog Post Deletions.” arXiv. http://arxiv.org/ftp/arxiv/papers/1303

Figure 1. Frequency of Geotagged Political Posts, Beijing, July 1, 2014−June 15, 2015

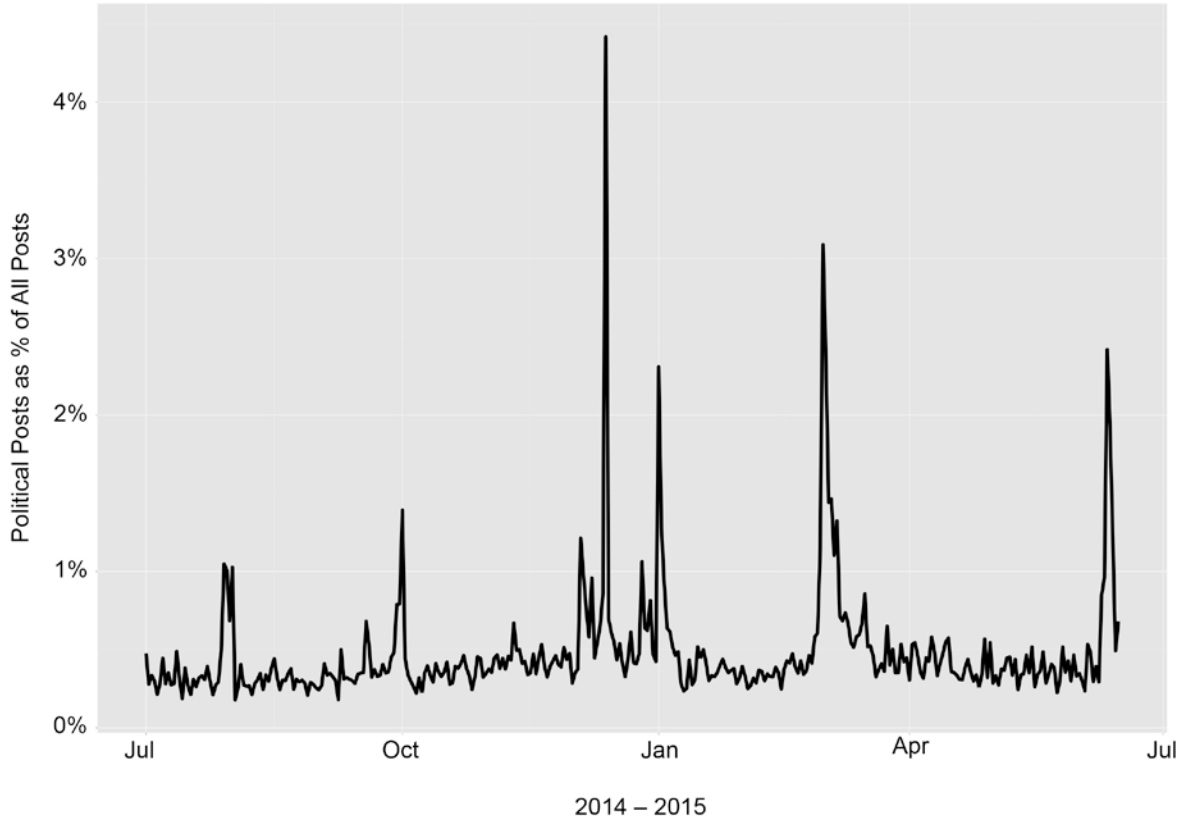


Figure 2. Spatial Distribution of Political Posts, Beijing, July 1, 2014−June 15, 2015

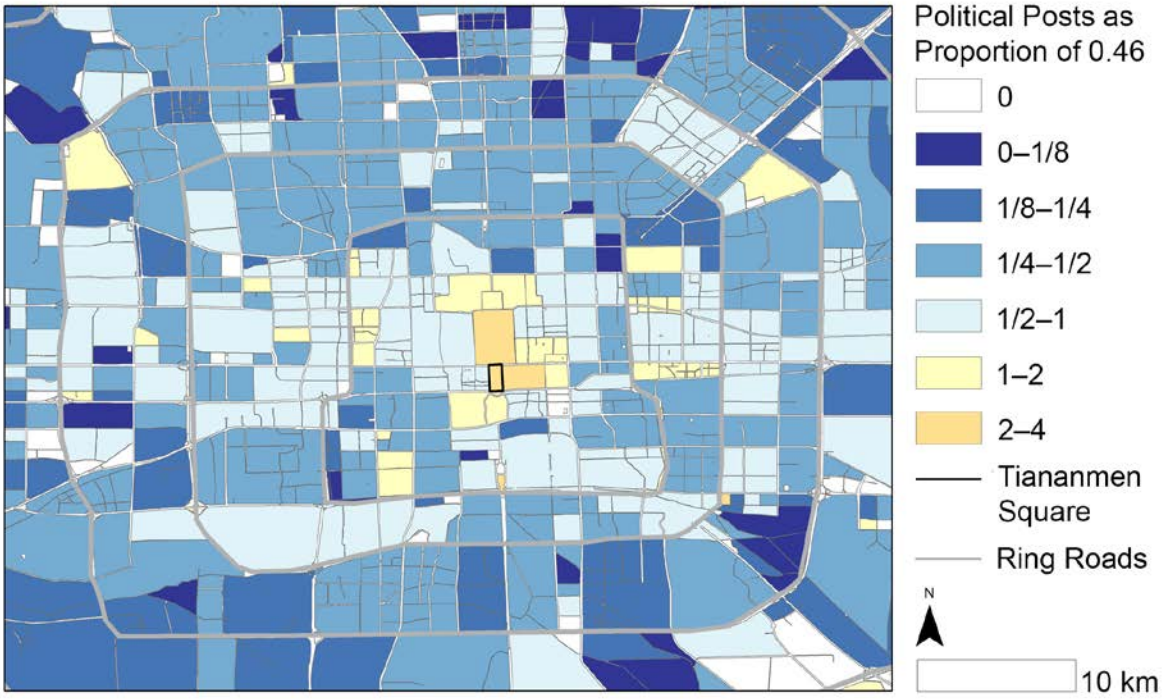


Table 1. Four Political Catalysts, 2014−2015

Political Catalyst	**Time 0**	**Determination of Time 0**
Mega-Tiger Zhou Investigation	18:00:00 July 29, 2014	Central Discipline Inspection Commission releases news of its investigation of Zhou Yongkang
Nanjing Massacre Memorial Day	07:00:00 December 13, 2014	Government lowers national flag and broadcasts this on morning news, official ceremony begins 3 hours later
Under the Dome	12:00:00 February 28, 2015	*Under the Dome* is released for mass viewing on internet
June 4	00:00:00 June 4, 2015	1989 anniversary day begins

Note: As *Under the Dome* was removed from all websites, the precise time of the documentary's release is unclear, but 12 noon is very close to the time.

Figure 3. Spatial Distribution of Political Posts, Beijing, Four Political Catalysts

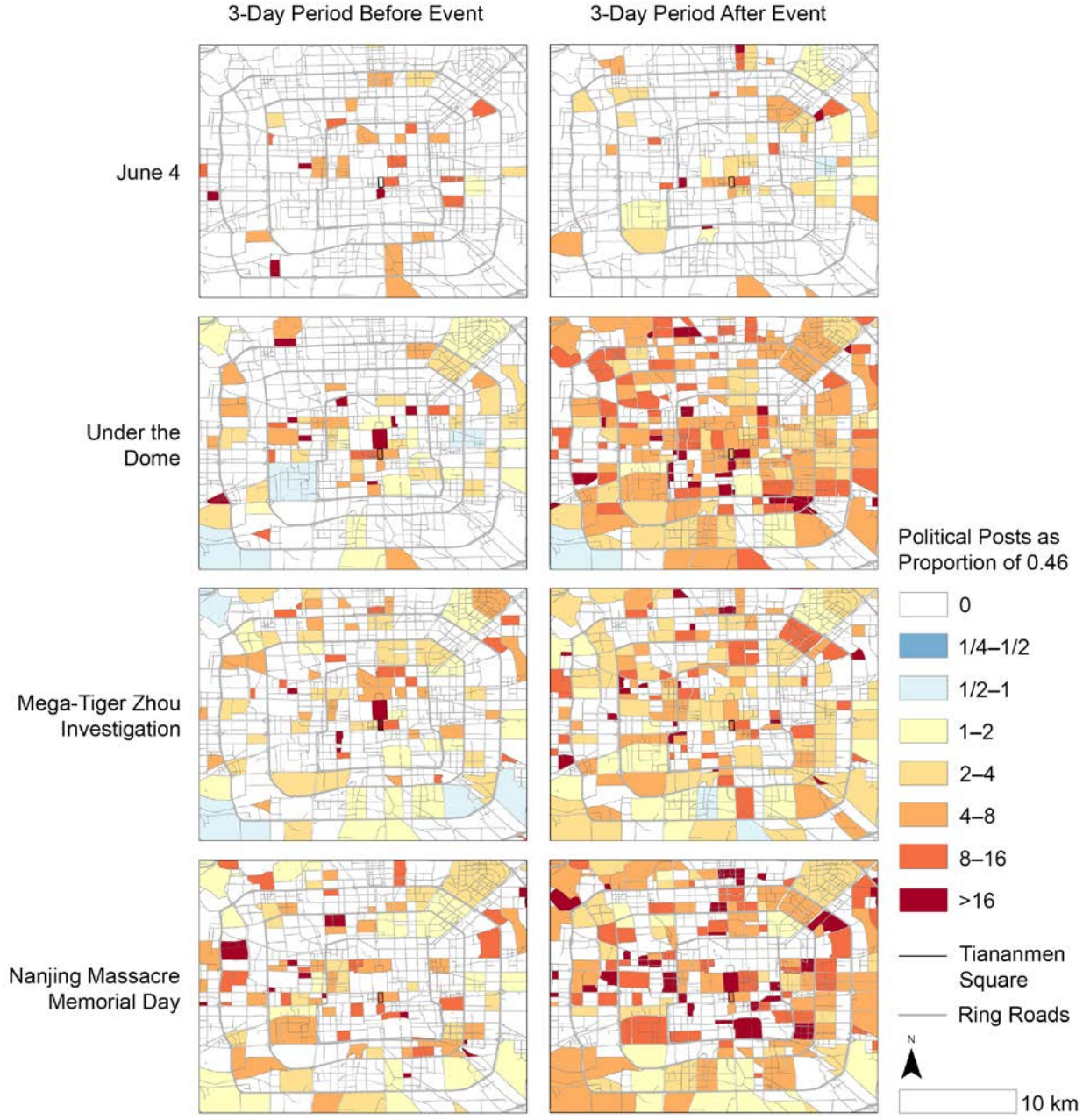


Figure 4. Spatial Distribution of Critical Political Posts, Beijing, Four Political Catalysts

Figure 5. Spatial Distribution of Critical Political Posts after Release of *Under the Dome*, Beijing

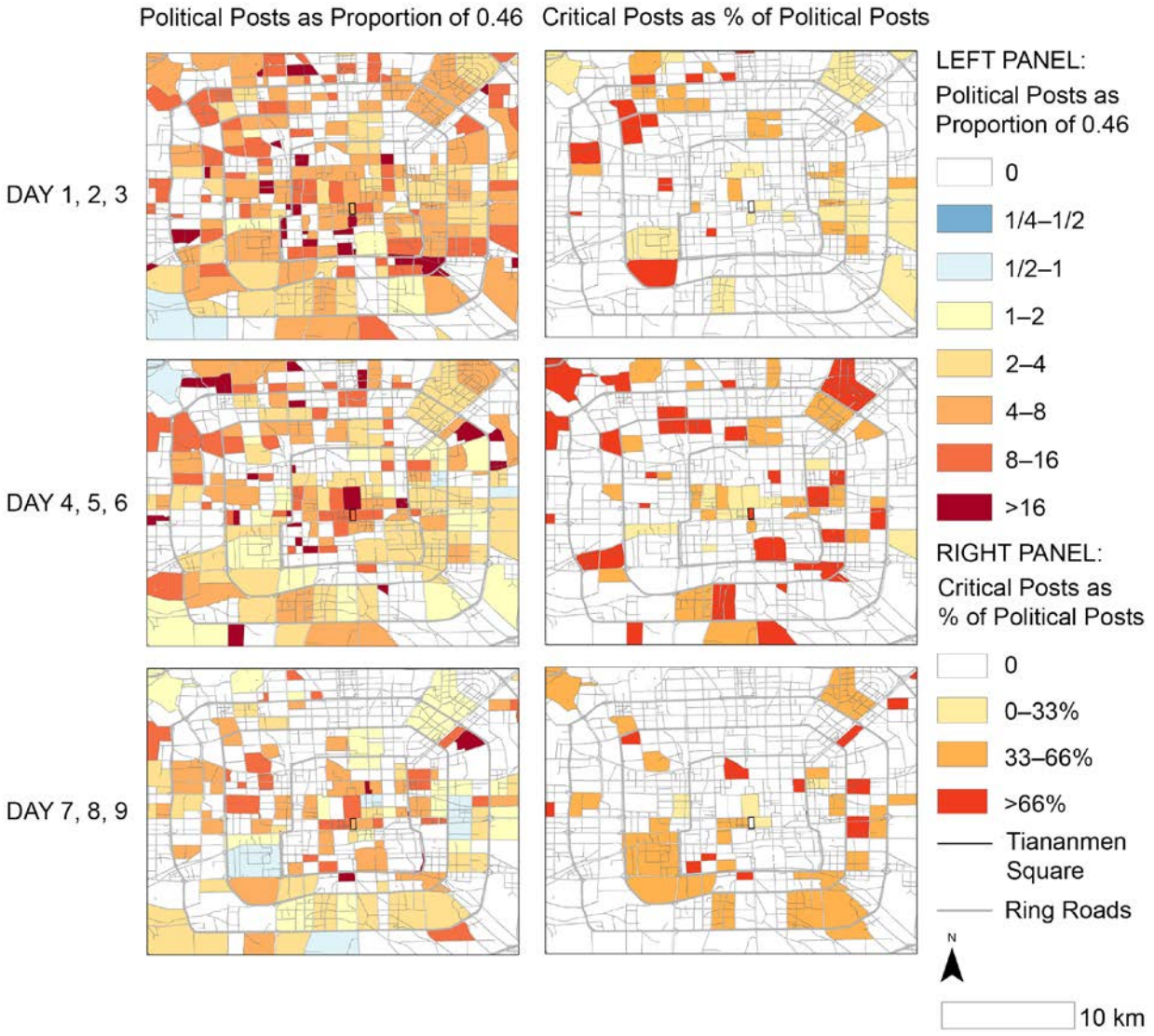
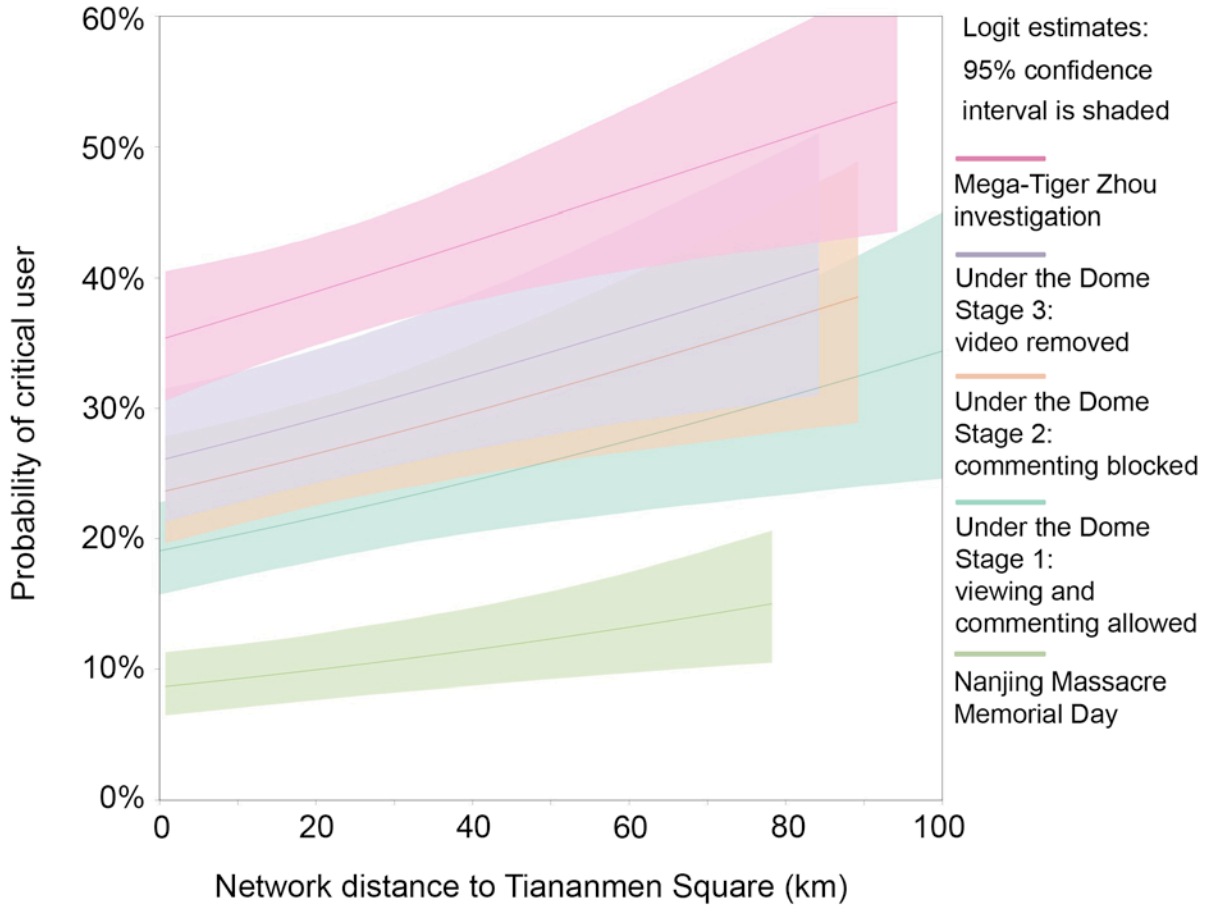


Table 2. Critical Posts and Distance to Tiananmen Square, Logistical Regression Models

	Model 1	**Model 2**
Network distance to Tiananmen Square, km	0.007**	0.007*
	(0.003)	(0.003)
0−72 hours: Mega-Tiger Zhou Investigation	0.924***	0.922***
	(0.14)	(0.14)
0−72 hours: Nanjing Massacre Memorial Day	−0.92***	−0.92***
	(0.18)	(0.18)
Stage 2, 72−144 hours: *Under the Dome* release	0.27	0.27
	(0.15)	(0.15)
Stage 3, 144−216 hours: *Under the Dome* release	0.13	0.13
	(0.17)	(0.17)
Sina Weibo followers (in 1,000s)		−0.00042
		(0.0011)
Sina Weibo followings (in 1,000s)		0.046
		(0.10)
Intercept	−1.43***	−1.45***
	(0.12)	(0.13)
Observations: 2,359 Sina Weibo users		
Dependent variable: critical post=1, other post=0		
Standard errors in parentheses, variable descriptions in text		
*p<.05, **p<.01, ***p<.001		

Figure 6. Critical Posts and Distance to Tiananmen Square



Appendix A. Data Collection

We collect and store geotagged posts from netizens in Beijing on Sina Weibo, for the 350-day period from 1 July 2014 through 15 June 2015. This amounts to more than 6 million unique posts. Geotagged posts are those sent on a device, typically a smartphone, with software enabled to identify the location of the device at the time of posting. The geotagged location is highly precise: the positional error is 2 to 25 meters (Zandbergen and Barbeau 2011). We are confident that we are collecting all censored geotagged posts in Beijing. The Sina Weibo Nearby Application Programing Interface (API) permits us, using a programming interface with customized parameters, to access the Sina Weibo database that stores all such posts. To systematically retrieve posts in Beijing, we designed a net of 95,480 location points set 400 meters apart from one another; this entirely encompasses the boundaries of Beijing municipality.[28] The Sina Weibo Nearby API limits us to retrieving from a location point, at any one time, a maximum of the 20,000 most recent posts.[29] We navigate our net of points, point by point, to retrieve the most recent posts within a radius we set at 1,000 meters. It takes us approximately three weeks to navigate every point in our net, retrieving posts at each point. By design, the radius of 1,000 meters for each of 95,480 location points set 400 meters apart from one another implies considerable spatial overlap, as illustrated below in Figure A.1. This contributes to intended redundancy in our retrieval of posts: in each three-week period, as we navigate our net of points systematically, from point to nearby point, at each point retrieving up

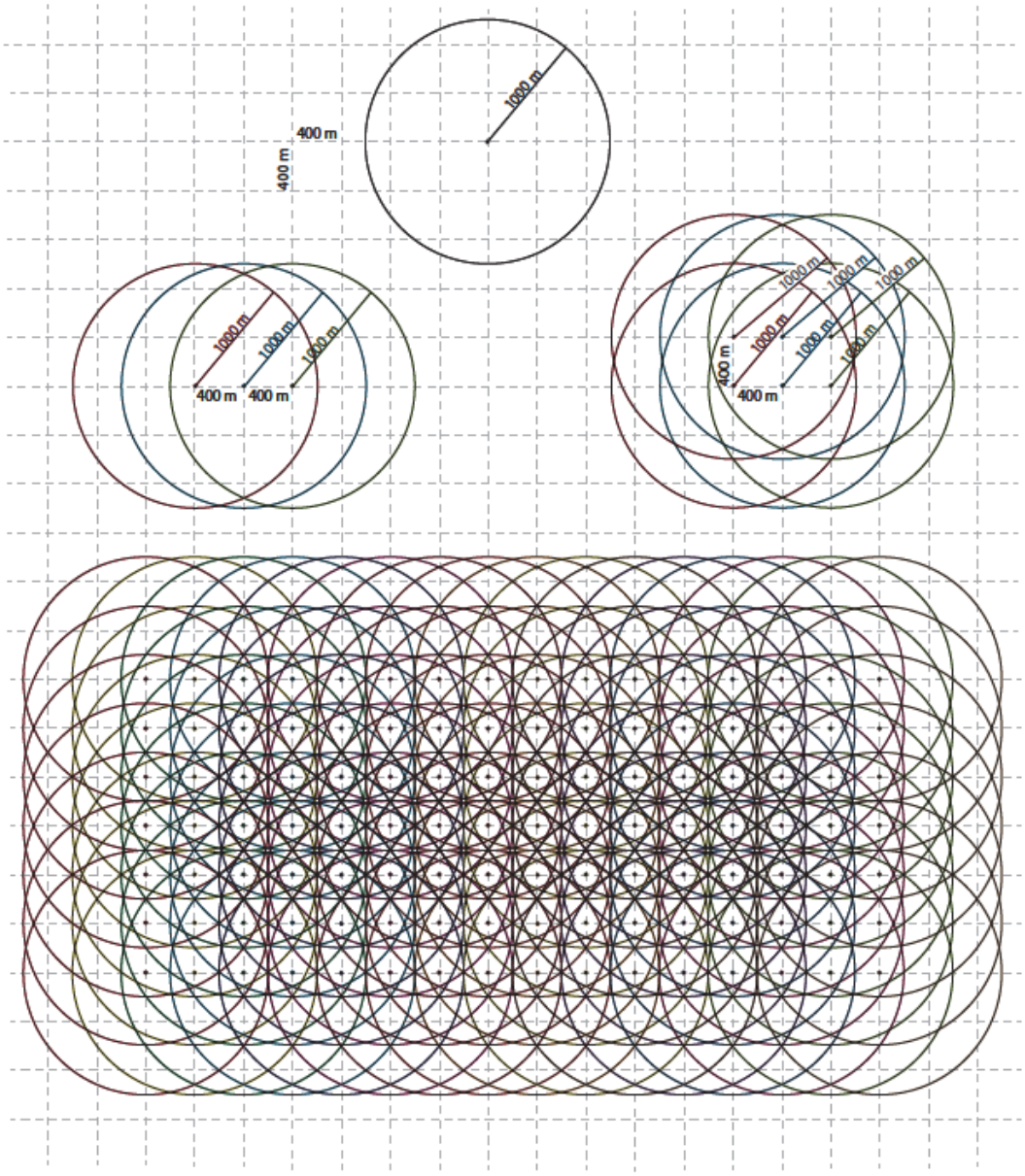
[28] It encompasses the area from 39.463 to 40.580 degrees north in latitude and from 115.815 to 117.176 degrees east in longitude.

[29] Specifically, Sina Weibo Nearby API has a technical limit of 50 posts per page and limits retrieval to the most recent 400 pages. If we view all 400 pages from a single location point at the same time, which we do here, we can view a maximum of the most recent 20,000 posts.

to 20,000 most recent posts within the 1,000-meter radius, our retrieval covers every piece of Beijing's geographic space many times over.[30] Given spatial and temporal overlap, with a net of 95,480 points located 400 meters apart and retrieval of all posts within a radius of 1,000 meters of each point every three weeks, we miss a post only if there are more than 20,000 posts from some single location in three weeks—which is effectively impossible.[31]

[30] We subsequently eliminate multiple observations of the same post to create a dataset of unique posts.
[31] Empirically, we find this limit of 20,000 posts in the retrieval space of a single location point is reached in about six months.

Figure A.1. A Net of Points in a Sea of Posts



Appendix B. Differences between Geotagged and Non-Geotagged Posts

Geotagged posts may be different from other posts, even intentionally different. If differences are significant, it threatens the validity of our research. Given our research question, we are mainly concerned about bias in the political content of geotagged posts. Are geotagged posts systematically more apolitical? Are geotagged posts with political content systematically less critical? We assess these potential threats to validity by making use of items in the BAS 2015 survey on use of geotagging and all questions about politics available to us. Specifically, we compare differences in political attitudes of netizens who report disabling the smartphone geotagging function when posting on Sina Weibo and netizens who report never having done so.

Among the 2,610 Beijing residents that the BAS 2015 surveyed, 1,609 (62 percent) are netizens. Thirty percent of them go online to access Sina Weibo—most of them at least several times a week, many of them several times a day. Smartphones are by far the most popular technology: 94 percent of BAS 2015 microbloggers go online by smartphone. We ask all netizens who have ever posted anything on Sina Weibo by smartphone if they know that smartphone posts can be geotagged: 92 percent of them say yes. Moreover, 39 percent of these geotag-aware microbloggers report they have disabled geotagging for at least some posts. How (if at all) do these microbloggers differ politically from those who report they have never disabled geotagging when posting on Sina Weibo? It seems plausible that those who have disabled geotagging are relatively more politicized, more politically sophisticated, perhaps more cynical about politicians and their official pronouncements.

To investigate this issue, we analyze differences between the two groups in terms of their interest in politics and views on the anticorruption campaign, as reflected in responses to five survey questions. Response frequencies are shown in Table B.1. We find no statistically

significant differences between those who report having turned off geotagging and those who report never having done so. Based on their responses, the two groups do not differ significantly in: their interest in politics, how much the anticorruption campaign affects their lives, their agreement with rhetoric about the campaign's purpose and the sincerity of efforts to fight corruption, or their opinion on whether or not the campaign should continue and intensify.[32]

[Table B.1 about here]

We also conduct a thorough search for other differences between the two groups. We find statistically significant differences on only two dimensions. First, microbloggers who have disabled geotagging are more highly educated: 64 percent of them have attended university, compared to the 53 percent of microbloggers who have never disabled geotagging.[33] Second, microbloggers who have disabled geotagging have more online followers: a reported 620 followers on average, compared to 280 for microbloggers who have never disabled geotagging.[34]

These results suggest that our focus on geotagged posts does not pose a threat to validity for our research. The two groups do not differ on the dimension of interest to us: political views. Moreover, the two differences of significance suggest that the netizens whose posts we analyze are more "ordinary" than very highly educated or more popular netizens. To the extent that this presents a bias, we find it acceptable. We purposely aim to shift attention from the political activists and celebrities that are the subject of most studies of the social media in authoritarian

[32] Specifically, the Pearson chi-square probability for observed differences in the distribution of responses across the two groups is higher than .05 (and usually much higher) for all five items. With this many observations, the Pearson chi-square test is very powerful, that is, highly likely to point to differences as statistically significant.
[33] This is not about the difference between university students and others, however. The difference in proportions currently studying, rather than working, is trivial and not statistically significant.
[34] For highest level of education, a categorical variable in the survey, we use a Pearson chi-square test; for number of online followers, we use a t-test of difference in group means. For both tests, we adopt .05 as our level of statistical significance.

settings to focus instead on more ordinary netizens.

Why have 39 percent of microbloggers at some time disabled geotagging when posting on Sina Weibo? We offer respondents a choice of eight reasons (and the opportunity to give another reason) for disabling geotagging. By far the most common choice, making up 39 percent of responses, is: “I don’t want strangers to know my location.” The second most common choice, accounting for 14 percent of responses, is: “I don’t want friends, family, or co-workers to know my location.” In short, in their online (inherently public) activity, Beijing netizens sometimes seek an element of privacy—not from Sina or the state, which is impossible in any case, but from other netizens.

Table B.1. Beijing Microbloggers Who Have (or Have Never) Disabled Geotagging, Percentages

	Geotag on	Geotag off
Interest in politics		
Very interested	3.9	5.4
Somewhat interested	28.9	27.0
Not too interested	59.5	62.2
Not interested at all	7.8	5.4
Impact of current campaign on your life		
Direct impact	11.5	11.0
Indirect impact	4.1	5.5
Basically no impact	84.4	83.5
Current campaign: highly important or high-level political struggle?		
Highly important	56.5	53.2
High-level political struggle	17.7	14.9
Both	19.1	29.1
Neither	6.7	2.8
Fighting corruption: response to mass public or political show?		
Reflects public sentiment, enjoys public support	64.7	60.3
Political show	7.0	5.9
Both	21.9	30.9
Neither	6.5	2.9
Approve of continuation and intensification of campaign?		
Very much approve	57.5	63.3
Somewhat approve	32.9	27.2
So-so	7.9	9.5
Somewhat disapprove	1.8	0
Very much disapprove	0	0

Notes: Percentages compare netizens who report having disabled the geotagging function for some posts when posting on Sina Weibo with netizens who report never having done so. Frequencies are responses to the following questions. (1) Are you interested in politics? Would you say you are very interested, somewhat interested, not too interested, or not interested at all? (2) Do you think the current anticorruption effort has an impact on your own life? A direct impact? An indirect impact? (3) Some view the current anticorruption effort as highly important: not to counter corruption may destroy the party and country. Others view it as a high-level political struggle. Which view do you agree with? (4) Some think fighting corruption reflects the public sentiment and enjoys broad popular support. Others think fighting corruption is a political show. Which view do you agree with? (5) Do you approve of the government continuing and intensifying the anticorruption effort?

Source: Beijing Area Study 2015

NETIZEN CONFIDENCE IN AUTOCRAT-DISPENSED INFORMATION OVER AN INTRUSIVE EVENT

1. Introduction

On Saturday, March 1, a group of knife-wielding Uighurs attacked the Kunming railway station. The state media quickly proclaimed this attack an act of "terrorism," and almost as quickly—only three hours later—it claimed the event "under control." Following these announcements, the state media, along with its commercial partners, rolled out a series of stories that dominated the entire media landscape for a few days. The stories were sensational: some were about heroic policemen risking their lives battling the terrorists; others told of the severely wounded who lived a nightmare they couldn't shake off. No doubt there were images of state officials warmly shaking the hands of the victims. As for civil society—celebrities among Uighurs, Tibetans, and athletes acknowledged by the state—all condemned the terrorists and vowed to stand by law and order. It would seem that Chinese autocrats, unlike politicians in a democratic country, were decisive and efficient in confronting terrorism and, by extension, any intrusive event. But is this anything like the true course of events?

Governments, democratic and authoritarian, have to deal with intrusive events. However, there is a stark difference between them in information transparency and political accountability. In democratic countries, media and social media that are independent from the political system disclose the factual developments of the intrusive event. And they are also dedicated to expose policy failures in relation to how government agencies and agents deal with the event. Society, through learning such information, can hold political leaders accountable in elections. In authoritarian regimes, however, dealing with intrusive events has one priority, which is to restore authoritarian rule as quickly as possible (Gandhi and Przeworski, 2007; Dunning, 2008;

Magaloni and Kricheli, 2010; Svolik, 2012).

In authoritarian regimes, information about intrusive events largely flows in a unidirectional, top-down manner (Huang, 2015). Society, for its part, tends to accept whatever information it is given, provided it is benign and conforms more or less to its own experience. On the other hand, because society knows that the regime has total control of the information, it can also be skeptical of it. When the skepticism becomes serious and widespread, the regime can and sometimes does make adjustments (Lorentzen, 2015). In China at least, the iron hand shows remarkable flexibility—this is how it maintains power!

To the high generalization above, I now offer a special case to understand in detail society's confidence in the information autocrats provide in the face of an intrusive event. That event is a terrorist attack on a densely populated area where eye-witnesses are likely to be the first on the scene, report what they find through the social media, and thus somewhat curtail the regime's ability to manipulate information for its own purpose. I focus on information manipulation of this intrusive event in cyberspace because it is there that the state media break the news and netizens first respond with their opinions. Social media on smartphones have in recent years become the gateway of information for Chinese society. Consider the fact that China's netizens comprise a population of 700 million by 2015, 89 percent of whom have access to the internet on smartphones (CNNIC, 2016). Empowered by such access, netizens are able to gain information about a sudden, highly stressful event from other netizens and sometimes even from Western media, in real-time, regardless of their whereabouts, before autocrats can apply

cyber-control. Given the enormous gain in communicative power,[35] isn't it plausible that netizens, triggered by an intrusive event, may show collective anger in cyberspace that poses a challenge to authoritarian rule?

This paper is organized as follows. Section 2 theorizes the power play between autocrats and netizens in cyberspace over the information that each contentiously puts forward. Section 3 introduces the hypotheses implied by the theory. Section 4 describes my empirical strategy. Section 5 is the empirical core of the article: it tests my hypotheses by analyzing netizen reactions to the information that autocrats put forward in cyberspace. Section 6 interprets the empirical findings and concludes.

2. Theoretical Framework

I theorize the contestation of information in nondemocratic countries and entities that occurs in cyberspace between two major players: autocrats and netizens (Figure 1). The information so contested is not specific facts about any policy, institution, or event; rather, it is about the degree that each player is willing to use information to advance its interest. Autocrats' interest is to sustain their authoritarian rule (Svolik, 2012). To forestall internal challenge, they monitor the social networks of their allies in government and army; and to forestall external challenge, they permit certain incendiary views to emerge online and monitor the early signal of discontent in society. As for netizens, they too advance their interests online. These include policies that concern security, environmental condition, and economic prospect; and institutional changes, which may be formal, as for example, the Hukou reform; or informal, as, for example, a

[35] Millions of Chinese netizens gained access to the internet two decades ago. Social media became popular seven years ago, and smartphones become widely used five years ago. Numerous advances in communication technologies have further made communication faster and at lower cost.

change in leadership that can increase autocrat retaliation against netizen political initiatives; or, to the contrary, a change that is of benefit to them. As for netizens, their trust of autocrats varies, depending on circumstance. In everyday life, they generally trust the information autocrats provide: even political dissenters do not question the announced procedures for obtaining permits, taking exams, and such like.[36] On the hand, even loyal followers of autocrats question some of the information they disclose.[37]

The players contest information in regard to its veracity and verifiability. Autocrats, small in number but rich in resources, can learn and verify netizen information through channels such as their internet surveillance program and through a large network of civil servants loyal to the state. Autocrats also have the advantage of being able to diffuse information widely, thanks to their control over the media. Since they lack netizen opinion towards their policies, they seek to gauge it by technological and planted-informer means. Netizens, for their part, have traditionally communicated through social circles, such communication being vastly expanded by technology. Netizens can use information to form political coalitions, a step that will put autocrats under pressure.

When autocrats' information on netizens is in excess of netizen information on autocrat (i.e., autocrats are confident of their ability to handle challenges that issue from what netizens know) and, moreover, when autocrats' information is non-communicative (i.e., autocrats' information does not spread widely), then autocrats are inclined to keep what they know hidden, though in so doing, they risk possible netizen distrust. However, when netizen information is

[36] O'Brien and Li have some excellent discussions on petitioners to some extent trust autocrats' rule and use them to their advantage in rural China.
[37] Some Maoists refuse to believe CCP's denouncement of Mao's fault in Cultural Revolution, despite they pledge their loyalty to the party.

superior to autocrats' information and is, moreover, potentially capable of spreading widely, autocrats are likely to disclose a certain amount of information to gain netizen trust.

[Figure 1 about here]

Intrusive events break the equilibrium between autocrats and netizens. Immediately after such an event, autocrats find themselves in the unusual position of having less information than netizens on the ground, who may well be eye-witnesses. Moreover, in an advanced communicative environment, netizens can spread what they know through their informal, subterranean social networks of friends and families. For this reason, information contestation between autocrats and netizens is tilted initially towards the right side of Figure 1. At this stage, autocrats may selectively disclose what they know to win netizens' trust, who for their part may reciprocate by offering autocrats, again selectively, information of their own. Gradually, the information contestation shifts towards the left side of Figure 1. As autocrats gain ground over netizens, they will hide what they know when the image it presents is unfavorable to them.

3.Hypotheses

Much study on information contestation between autocrats and netizens have concentrated on the authoritarian control of topical contents in media and social media (Egorov et al. 2009, Brady 2008, 2009, Hassid 2008, Kennedy 2009, Hong 2011, Stockmann and Gallagher 2011, Tang and Sampson 2012, Gehlbach et al, 2015). Content in media is a measure of what readers want to read and a measure of what autocrats permit that is unfavorable to them. Such measures cannot, however, do justice to the contestations because autocrats control the topical contents and the political reactions from netizens are seldom, if ever, known. In authoritarian China, scholars find that incendiary nationalistic stories can ignite a permanent ground of discontent in the people, inciting them to strong criticism and protest (Stockmann

2013, Cairns and Carlson 2016). As for political activists, their venue of criticism and protest is the social media (Lei, 2011).

The rise of social media offers us an opportunity to gauge information contestation with greater clarity (Morozov 2012, Kelly et al., 2012, Greitens 2013, Lorentzen, 2014, Noesselt 2014, Gunitsky 2015, Enikolopov et al, 2016). Studies show that political talks on social media over a sudden, disrupting event has been removed or diluted to prevent collective actions (King, Pan, Roberts 2013, 2014, 2016). These studies focus, however, on the opinion of political activists, who are few in number, or they are based on an experimental design. In either case, a sudden, disrupting event is found to cause contestation between autocrats and netizens. My study complements these works in that it takes into consideration, not just political activists, but the far larger population of netizens that may, of course, include political activists. Moreover, my study is based on a real-world sudden event, which makes it a "natural" experiment, enabling me to gauge the complexities of information contestation between autocrats and netizens in greater detail than was hitherto achievable.

The sudden event is the attack at Kunming Railway Station. Kunming is not a very large city by contemporary Chinese standard, nor is the number killed—thirty-one—an unconscionably large number. Rather it is the impact of such attacks on overall social stability that matters. Also of importance, certainly to autocrats themselves, is their own mandate to rule. Both depend not only on facts on the ground but also on how they are communicated. Autocrats are well aware of this. Thus, in the case of the Kunming attack, they sought to dominate information in their engagement with netizens by swiftly breaking the news. Moreover, they provided updates from various media sources and did so with such frequency that it was as though the event was broadcast by live streaming. Minutes following the attack, the media

reported the first story: "A group of criminals burst into the Kunming Railway Station and stabbed whomever they saw (story with pictures). "[38] At 11:27 pm, only two hours after the event, the media reported that the Kunming Railway Station had resumed operation. According to *People's Daily*, at 1:30 am on 2 March 2014, on Xi Jinping's order, Politburo member Meng Jianzhu boarded an airplane to Kunming. At 3:45 am, the party's official news agency, Xinhua, released a statement from the Chinese president. At 7:11 am, Xinhua announced that the attack had been carried out by Xinjiang separatists.

As for netizens, their reactions to such information disclosure may be analyzed as follows. Netizen response will present a problem to autocrats when they are either not engaged by the information or they reject it publicly. When netizens reject it publicly, autocrats may be obliged to censor the rejection or to replace it with fake comments. Failure to engage netizens at all is a serious matter and will prompt autocrats to make policy and institutional changes or to purge local politicians. On the other hand, if netizens *are* engaged by autocrats' information, autocrats will have the advantage of framing the discourse. I hypothesize the empirical test as follows.

H1: Netizens are engaged by the information that autocrats disclose about the intrusive event.

In this hypothesis, I examine the netizens' response in chronological order of the official announcements following the attack at Kunming Railway Station. Specifically, I examine how netizens, whether political or non-political, engage with the state media's breaking news. What

[38] According to the timeline released by the state media, the attack took place at approximately 9:20 pm and lasted 25 minutes. Sina first broke the news at 10:22 pm on its website, after its corresponding Weibo accounts broke the news on Weibo.

are their reactions both before the report and after it until the end of the event is officially announced? My hypothesis is that when such a disaster occurs in an authoritarian regime, autocrats, using their media resources, proceed to "sell" netizens their own version of what took place.

Autocrats' control of media and other political resources gives them a decisive image-making and story-telling advantage, and this holds when an intrusive event occurs. One challenge is simply that of giving shape and order (or frame) to the chaos of happening. A frame endows the framer with an air of authority—of being in control (Goffman, 1974). From the autocrats' point of view, this step is necessary to forestall the quick spreading of unauthorized, subterranean information. Once the frame is in place, netizens have to decide whether they will buy it. To the extent that they do buy it, they will seem to follow the autocrats' lead.

My first hypothesis is designed to test the immediate reactions of netizens to the information that autocrats put forward. As I noted above, if netizens dismiss the information or confront stiff censorship, autocrats will find themselves under pressure either to make further changes in policy to keep netizens content or face strong collective protest. However, if netizens are engaged by the information, further examination is required to understand whether such engagement constitutes trust, for Chinese netizens can be skeptical and even cynical in regard to what autocrats—indeed, what all authorities—tell them. This leads to my second hypothesis.

H2: Chinese netizens are not only engaged by but also trust the information about the intrusive event.

Having posed the question of netizen trust in the information that autocrats disclose, I now extend my enquiry of such trust to the socio-political background of the netizens. My question is: does netizen trust in the information disclosed by autocrats differ by socio-political

background? A growing literature shows that Chinese autocrats are responsive to netizen concerns and demands, a responsiveness that should earn netizen confidence (Meng, Pan, Yang 2014; Chen, Pan, Xu, 2015; Manion, 2015). I, however, explore the limits of such confidence even with netizens who are the beneficiaries of authoritarian rule. In particular, I offer two hypotheses (H3 and H4) to the effect that netizen trust varies with wealth and power.

H3: After the intrusive event, netizens with connections to government are more prone to trust the information given by autocrats.

H4: After the intrusive event, rich netizens are more prone to trust the information given by autocrats.

Netizens connected to government, either as government officials themselves or as families or friends of officials, have more access and exposure to government information; they should be more likely to trust the information autocrats provide after the intrusive event. Also, netizens with prestigious socioeconomic background, having benefited from the autocrats, are more likely to trust the information that autocrats provide.

4. Empirical Strategy

To test my hypotheses, I turn to Chinese netizens in cyberspace. In recent years, thanks to the rising popularity of smartphones, Chinese netizens are able to gain information about sudden events with unprecedented speed. As reported above, China has around 700 million netizens by the end of 2015, of whom 620 million use smartphones (CNNIC, 2016). Of the new netizens, a majority primarily or purely use smartphones to access the internet. Moreover, a growing number of Chinese use news applications to register their preferences, which they do either by the number of hits or by comments, and post their opinions on social media. Such digital traces

enable me to understand subtle and detailed behaviors, including the opinion of individuals as well as the extent and time of their personal trips.

Digital traces from smartphones offer me empirical evidence for the Kunming Railway Station attack. Their source is Sina, perhaps the largest media corporation in China. Sina provides a unique public forum for Chinese netizens. It features integrated smartphone media and social media applications that attract hundreds of millions of users.[39] More importantly, accessibility through search engines by anyone is pivotal to all Sina's services, in contrast to conglomerates such as WeChat that concentrate more on facilitating private communication. Given its open, public nature, Sina News and Weibo offer a bridge between the authoritarian state and society, providing a rare opportunity for the state to engage netizens and for netizens to make demands on the state.

The literature suggests the limits of commercial media such as Sina in reporting intrusive events. Stockmann (2013) distinguishes between open news topics (on which the media can report quite freely) and closed (mostly political) topics, subject to censorship and self-censorship. The Kunming Railway Station attack, which the party defines as "an attack initiated by Uighur separatists," is a prime example of a closed news topic on which the commercial media justify coverage with reference to "rule of the market." Sina News and Weibo share with each other a defining feature: user-generated content. The user-generated content, including views, comments, reposts, and original posts, is recorded in a Sina account for each user, which can be used to log in both services with her smartphone. Sina News and Sina Weibo can report the story

[39] CNNIC reports that Sina Weibo occupies 70% of microblogging services, so that Weibo in Chinese virtually stands for Sina Weibo.

simultaneously, reaching a broad audience through smartphones.[40] At the same time, Chinese state intrusiveness on both Sina News and Weibo reflects its overwhelming obsession with social stability: communication about collective action is highly censurable (King, Pan, and Roberts 2013).

The Kunming Railway Station attack highlights Sina's role in presenting a more or less balanced account of news that satisfies the different requirements of state and society. Both Sina News and Weibo covered the intrusive event fully. Sina News, taking advantage of early netizen responses on Sina Weibo, immediately launched a special rolling report, one that broke 80 stories in 24 hours. Stories that Sina News and Weibo promote are mostly edited or closely monitored by government officials and journalists, who undoubtedly follow the official storyline (Stockmann 2010, 2013; Nip and Fu 2016), but they also allowed netizens to express sympathy and rally behind the state in the form of comments and posts (Chang and Manion 2016).

For both autocrats and netizens, their strategic choice of information in cyberspace is registered in Sina News and Sina Weibo, both before the attack and immediately after it. It is the immediacy and spontaneity of the recording that makes information in cyberspace different from information usually obtained in interviews and surveys, which may be days, weeks, months, and even years later and is in this sense remembered, historical, and almost certainly somewhat altered.

To test the four hypotheses that I proposed in the previous section, I investigate the content of news stories pertinent to autocrats on Sina News, as well as the content and geo-

[40] The smartphone Apps of Sina News and Weibo are available for different smartphone platforms, such as iOS or Android, and have many features that are similar to USA Today App and Twitter App respectively.

location of news comments and social media posts pertinent to netizen response on Sina Weibo both before and after the attack. I gauge autocrats' decisions and, most importantly, netizen confidence in autocrats' information in a direct manner, one that has not been previously tried in China.

4.1 Data Collection: Stories and Comments from Sina News and Geotagged Posts from Sina Weibo

I collect political stories that appear in Sina News and netizen comments on these stories, including stories about the intrusive event and geotagged Weibo posts sent from Kunming netizens as these appear before and after the event. I further use the geolocations that netizens have made public while posting on social media to examine their movements around two focal places—Kunming Railway Station and the Dashuying Uighur community—both of which were reported to have been attacked.

Why are Sina News and Sina Weibo my sources? Sina is a leading Chinese commercial media company, incorporated in 1999 and listed on the NASDAQ since 2000. Its success is built on attracting an ever-bigger readership so as to increase revenues, which derive primarily from advertising. Success depends on more than audience appeal, however. News coverage also reflects the constraints of Chinese politics. In attending to the bottom line, Sina executives must correctly interpret the nuances of the politically permissible, because they bear ultimate responsibility for the content and consequences of whatever appears on the company's website. Among commercial media companies, Sina counts as a very compliant partner of the Chinese state (Cairns 11 August 2015).

Sina News is one of China's largest news websites. It publicizes hundreds of news stories a day. Its stories originate from a variety of sources, including traditional newspapers, news

magazines, government websites, smaller news websites, independent columnists, and bloggers. The stories are subsequently edited by Sina and placed on its website and smartphone apps (applications) with a layout that has the look of a daily newspaper, with important news stories given top billing in the headlines. Exciting stories, decorated with photographs, are near the top of the page. All stories, important and mundane, are itemized on the page as shortened headlines. Netizens click or tap on headlines (or images) associated with stories to view the stories in full on separate pages.

Integrated with Sina News is Sina Weibo, the smartphone app that has become a part of such everyday experience for Chinese netizens, and is endorsed by the Chinese state.[41] Sina Weibo is a social media specialized in microblogging (Twitter-like 140-character blog) services and has become popular since 2009. Moreover, because a single Chinese character can constitute a complete word, the brevity requirement for posts is not so constraining. As artist activist Ai Weiwei (2011, 241) puts it: "In Chinese, 140 characters is a novella."

Sina Weibo so dominates microblogging services in China (CNNIC, 2015) that weibo (微博, Chinese for microblog) effectively means Sina Weibo. And because of WeChat, the popular instant messaging application on smartphones, Sina Weibo users can share information on Weibo privately to their friends who may not be in the circle of followers. Anything expressed on the platform can be publicly, rapidly, broadly—in short, virally—transmitted through reposting and private sharing. A Chinese Academy of Social Sciences study (Xinhua 24

[41] Sina Weibo, along with Tencent Wechat, is promoted as "两微一端" (i.e., Two 'W's on smartphones), which is endorsed by the highest administrative apparatus—the Central Leading Group for Cyberspace affairs and Cyberspace administration. (http://www.cac.gov.cn/2016-03/16/c_1118344118.htm)

July 2015) finds Sina Weibo the source of most "rumors": 59 percent, compared to 2 percent for traditional social media.

As Table 1 shows, my analysis consists of 13,103 stories about Chinese politics on Sina News from 1 November 2013 to 6 March 2014: 12,969 of them in the five months before the deadly event and 134 stories about the attack in five days after the attack,[42] all of which can also be accessed by Sina News apps on smartphones. I focus on netizens who are physically present in Kunming because they are the group of netizens that contest autocrats' information concerning their security. In addition, local people, including their commercial press, are almost certain to get to the scene first and what they see and report are quickly diffused by contemporary social media. That they are physically present I can tell from their use of smartphones. I can also tell that these smartphone users are private individuals, not journalists, political activists, or paid commentators. According to Sina, of 5 billion comments that are elicited by these stories, only approximately 70,000 of them are shown publicly and accessible by ordinary viewers. Each comment is also associated with a rich meta-data that includes the user's name, id (if available), IP address, the type of recording-and-sending device, the type of web browser, and the send-time. I analyze comments on the Railway Station attack from 375 Kunming netizens who have both smartphone and Weibo id.

Using Sina Weibo Nearby API, I also collect 1.45 million social media posts with precise geolocations from 1 November 2013 to 1 August 2014 (Appendix A). The posts so collected are sent from the smartphones and tablets of 305,643 netizens; they are therefore most unlikely to be

[42] Stories come from the section of Chinese domestic news (国内新闻), the section about Chinese politics. Stories about society and international relations have their own sections.

fabricated or distorted by propaganda or social media bots (Appendix B). Sina Weibo posts, like those of Sina News, also carry rich meta-data generated automatically by smartphone Weibo App. These include sent time, device type, and sent location, as well as the netizen's number of followers and the cumulative number of posts at a particular moment. The geolocation that comes from the GPS module of netizens' smartphones is the precise location (i.e., 2-10 meter accuracy) of netizen whereabouts at the time the posts are sent. This precise location is known only to developers, Sina, and the State, but not to other netizens. In a survey conducted with 2,620 respondents concerning their use of smartphones and social media, I obtained results that validate the similarity between geotagged and non-geotagged Weibo posts (Appendix B).

I analyze geotagged posts sent from Kunming for three reasons. First, they are sent from smartphones and tablets, and carry personal opinions. As suggested by another study, the smartphone and tablet users who sent these posts are likely to be authentic netizens who are truly interested in news and politics, instead of celebrities, journalists, paid-commentators, advertising bots, and political activists (Fu, Samet, and Sankaranarayanan, 2014). Second, the revealed locations confirm that they are indeed netizens in China, not political activists overseas. According to another study, netizens in China are not so outspoken in politics than those out of China, especially after the real-name registration policy was adopted (Fu, Chan, and Chau, 2013). Third, the geo-locations associated with posts are least likely to be manipulated, in comparison with the numerous measures of information manipulation on Sina News comment section and Sina Weibo (Appendix C).

4.2 Testing Strategy for Hypothesis 1: Comparison of Netizens in Sina News Comment Section and Sina Weibo

I integrate the Sina News comments and Sina Weibo posts by linking user's Weibo id that I identify from the 305,643 netizens who sent at least one geotagged post during the ten-month study period. 375 netizens are found reading Sina News on their smartphones and occasionally commenting on the stories, and I use these netizens to test Hypothesis 1. I define the 5 days *after* the attack as the attack information period, a period that corresponds with the news cycle of Sina News when the Kunming Railway Station came under attack and when the event elicited widespread discussion. I compare netizens who comment on Sina News stories in that period with those who comment on "normal" topics five months earlier.[43] We know that the authorities manipulate information emanating from Sina News. However, in the test, I assume that such manipulation is independent of the comments of authentic Kunming netizens on stories on Chinese politics. Assuming that the comments of netizens are consistent over time, my test of comparing netizens in the 5-month period before the attack with those in the 5-day period of the attack is a conservative test: if netizens before the attack are significantly fewer than those during the period of attack, it provides strong evidence that authentic netizens are engaged by the stories that autocrats disclose.

I examine netizens who sent geotagged posts in Kunming on Sina Weibo as my second set of empirical evidence. I compare netizens who post on Sina Weibo with geotag in the attack information period with those who post with geotag five months earlier. Netizen posting on Sina Weibo with geotag is customary, as my survey suggests (Appendix B). Therefore, netizens posting with geotag before and during the period of the attack should be similar, unless their behavior has changed abruptly following the attack. If netizens post with geotag before the attack

[43] Five days defined here start at 9pm on 1 March 2014 and end at 9pm on 5 March 2014.

are significantly fewer than those during the period of the attack, it will be evidence that authentic netizens are engaged by the stories that autocrats disclose.

I further identify political netizens in this test, because they, by virtue of being especially sensitive to political pressure, may post differently. To identify political netizens, I analyze the digital traces left by them—taking note of whether their posts on Sina Weibo have political content and whether they comment on stories about Chinese politics as these are reported in Sina News. To distinguish political posts from the 1.45 million social media posts, I assemble 522 Chinese characters that are highly related to politics, and use them to detect whether the social media contains any of them. For example, the character 'party' (党 *dang*) is politically charged. Therefore, any post that contains words such as 'party member' (党员 *dang yuan*) or 'anti-party' (反党 *fan dang*) or other combinations of words with 'party' in them are treated as potentially political. Because political posts only constitute less than 1% of Weibo posts,[44] advanced text classification and topic modeling can disproportionally omit them. My method, by contrast, is more accurate in the inclusion of political posts, and this is true even though it may also allow a large number of irrelevant ones to enter.[45] I then code through visual perception 19,852 political posts from this sample. I divide the 305,643 netizens in my study period into two groups—political and non-political. In the political group, I identify 13,995 netizens who sent in a total of 19,542 political posts at various times. In the non-political group, I identify 291,648 netizens. I then use the political netizens as the control in my test.

4.3 Testing Strategy for Hypotheses 2, 3 and 4: Space and Place in Netizen Action

[44] My visual coding on more than 30,000 Weibo posts from Beijing finds that only 0.4% of posts are political.
[45] Using an independent sample of 3,977 political posts on Weibo for validation, my method successfully captures 95.6% of these posts, achieving an omission error of 4.3%.

Having identified a strategy to test whether netizens are engaged by the information disclosed by autocrats, I now turn to examine netizen trust in such information. Netizens engaged by the information that autocrats put forward is different from trusting the information in authoritarian regimes. Following the intrusive event, social coercion and grief can be oppressive, in parallel with the monotonic storyline propagated by autocrats. Lacking a different narrative from a verified source other than state-run media, netizens who have different opinions cannot express it publicly, for fear of being judged heartless and inhumane. Public opinion of netizens during the period of the attack can be a "public lie." (Kuran, 1997) It is plausible that netizens provisionally or ostensibly accept the information that autocrats disclose, either in fear of autocrats or in fear of societal disapproval, or both. If so, the content of netizen posts on Sina Weibo misrepresents what netizens really think.

To investigate whether netizen engagement is consistent with netizen trust in autocrats' information, I devise an original and unobtrusive measure of netizen trust. Private actions, unlike public opinions, are more likely to reflect netizen trust on autocrats' information authentically. Because autocrats quickly proclaim that Kunming is safe—especially Kunming Railway Station and the Dashuying Uighur Community—I examine whether netizens show a curiosity, instead of concern with their safety, by frequenting these places. Given that location at the time of posting is a part of metadata only visible to Sina Weibo and its developers, its change can be viewed as a variation caused only by the intrusive event, equivalent to a natural experiment. Using netizen locations both five months before and five months after the deadly attack, I examine whether netizens return en masse to the Kunming Railway Station and the Dashuying Uighur Community after the attack, which indicates high trust in the autocrats' safety claim; or netizen avoidance of the Kunming Railway Station and the Dashuying Uighur Community after the attack, which

indicates a distrust of autocrats' safety claim; or netizens do not behave any differently after the attack, which I take as an evidence of a lukewarm acceptance of autocrats' safety claim.

I identify 57,081 Kunming smartphone users who post on Weibo with their geolocations before and after the attack. Some of them post multiple times, producing a total number of 778,922 posts. Given the locations they reveal with their posts, I can estimate the road distance from where netizens post on Weibo to the Kunming Railway Station and the Dashuying Uighur Community by means of a GIS-based network analysis and a sophisticated, "turn-by-turn" system of GPS navigation.[46] The approximately 778,922 locations recorded in netizen instruments allow me to run two regressions with user fixed effects, using, respectively, the distance to the Kunming Railway Station and to the Dashuying Uighur community before and after the attack.

The locational change of netizens before and after the attack give me further insight into netizen trust for Hypothesis 3 and 4. This is because where they live and work reveals their socio-political background and status. Individuals are not allowed to enter government compounds and, for that matter, even gated communities unless they have the proper credentials, for example. To proceed, I take advantage of the results obtained in the first paper of the dissertation. Specifically, I use government POIs to identify the offices and residences of government officials and their affiliates, mostly exclusive to those at provincial and municipal level (Appendix D). Netizens who have sent at least one post from such governmental entities during the ten-month study period are taken to be government-related netizens. Those who have

[46] Baidu direction API is used, which provides GPS navigation for driving, walking, biking, and public transit. I estimate the distance using the driving (or taking taxi) mode here.

never done so are taken to be non-governmental netizens. In a similar way, I identify netizens who send posts mostly from gated communities as well-to-do residents. I then use a difference-in-differences model to capture the potential effects of socio-political background of netizens (Earle and Gehlbach, 2015). In the regression, the interaction term of post-attack and appearance in the exclusive space such as governmental entities and gated communities captures the differential effects of the Kunming Railway Station Attack on netizens from different socio-political backgrounds. Should the estimate of this interaction term be significantly positive, it means the group of netizens in the interaction term has moved significantly farther away from the focal places—an indication of distrust in the information disclosure from autocrats. Otherwise, it indicates moving closer to the focal places than other netizen groups and so shows that group's trust in the information disclosure.

5. Findings

5.1 Hypothesis 1 Supported: Real netizens in Kunming respond to the information that autocrats disclose about the intrusive event.

Sina News quickly publishes stories from various state sources, which is an indication that Chinese autocrats have decided to give full information disclosure in accordance with their own needs and purposes. On a rolling basis, 80 stories were published in the first 24 hours, and altogether 134 stories were published in the total five days. Most of these stories stayed in headlines, eliciting millions of netizen hits during the five days after the attack. The stories mostly come from official sources—78 of 134 stories come from the state-run media and government websites, and the remaining from commercialized media, which suggests that the news stories about this intrusive event are closely overseen by Chinese autocrats.

Information manipulation by autocrats is simple and clear: emphasis is on how quickly the perpetrators are named—they are the Xinjiang separatists—and how quickly the attack is contained. Sina News stories, taking full advantage of the resources of Sina Weibo, even try to sensationalize the event with gaudy pictures. As the official storyline expands in scope, censorable information may appear as, for example, a story in Sina News titled "Dashuying Uighur Community is also under attack." It was quickly deleted.

[Table 1 about here]

Does the attempt by autocrats to capture the attention of netizens work? Evidence from Sina News comments suggests that it does. In Kunming, 295 smartphone users comment on the 134 stories generated by the event, as compared to 80 users commenting on all other Sina News stories in the five-month period before the attack (Table 1). Frequent official stories of the attack have, however, the effect of making Kunming netizens feel insecure. They demand tougher rule. Their comments show anger, fear, and suspicion toward minority groups. They do not, however accuse the party of wrong-doing, nor do they call for policy changes. Below is a sample of the more typical comments:

- All behaviors that injure the ordinary people should be punished!
- This is not an ethnic issue, not a religious issue, but a life-and-death struggle about a safeguarded national unity, social stability and people's happiness!
- They [the suspects] must be drug dealers.
- How could this be? Bless Kunming…

The fact that the number of authentic Kunming netizens who comment on news of the event far exceeds the number if authentic netizens who comment on political news generally before the event supports my Hypothesis 1, namely, authentic netizens are engaged with the

information disclosed by autocrats and are generally supportive of it. To be sure, the total number of netizens who comment on Sina News I identify as smartphone Sina News users in Kunming is small and subject to distortion. In addition, I find that they who posted on Weibo in regard to the attack greatly outnumber those smartphone users who commented on political news generally before the attack, and this further backs my Hypothesis 1. It should also be noted that posts are sent from every community in Kunming and the volume is astronomical (Figure 2). Moreover, everything from Sina News stories to comments inspired by them to Sina Weibo posts demonstrate that the Chinese authoritarian regime has such a command on cyberspace messaging that it can take control over the reporting of almost any intrusive event. When it occurs, commercial and social media may be the first to break the news, but the regime can quickly co-op it with its own take on what has happened.

[Figure 2 about here]

To answer the question of how autocrats manipulate netizen opinion, I compare netizens who have and netizens who have not voiced political opinions in the five months prior to the attack. The comparison again supports Hypothesis 1 and offers, moreover, some explanation of such outsize support (Table 2): 83 percent of the 2,412 netizens who were political in the period of the attack were non-political before the attack. Only 27 of the 1,540 political netizens who were political in the five months prior to the event also posted political opinions during the event. The findings show that most netizens who were political in the five-month period before the event either chose to post opinions that were non-political or chose not to post at all. However, a large number of previously non-political netizens—I call them “political outsiders”—suddenly turned political, taking up positions that show trust in the regime and patriotism for the country.

[Table 2 about here]

5.2 Hypothesis 2 Rejected: Chinese netizens, though they may be engaged with the information concerning the intrusive event, do not (on average) trust it.

I now turn to an indirect measure of trust in information, that measure being netizen movements in and around focal places reported as having been attacked or were subject to attack. Thanks to social media posts, I am able to map netizen movements at two time markers, one after the event and the other five months before it. The latter I deem to be the "norm" which gives me a basis for comparison and to test my second hypothesis.

Let me first describe the two critically important places of the crisis: the railway station and the Uighur community of Dashuying. The railway station is located between Huangcheng South Road, the inner most ring road of Kunming, and the south Second Ring Road. The Uighur community, Dashuying, is located next to the east Second Ring Road. Both places are located in the populous downtown area and are frequent stops for people to shop, do business, or change buses. That's where the terrorists struck—the sensitive urban core. Moreover, the Chinese authoritarian government quickly sought to reassure the people, claiming to have re-established order in a matter of hours. It also tried to deflect blame from the Uighur community in Kunming and from the Uighur people as a whole.

If netizens trust government information, they will quickly return to their habitual haunts, or they will approach the attacked localities, now proclaimed safe, and to do so out of curiosity. Evidence suggests that neither recourse happened. After the attack, the 57,081 netizens who posted with geotag before and after the event significantly moved away from both Kunming Railway Station and the Uighur community: on average, 0.65 kilometers away from the former and 0.68 kilometers away from the latter (Model 1 in Tables 3 and 4). These distances take

approximately ten minutes to cover by foot and are considered "safe" if the area of potential crime is densely peopled and the weapon used is a simple knife (Figure 3).

[Tables 3, 4 and Figure 3 about here]

The findings reject Hypothesis 2: Chinese netizens, ever concerned with security, avoid places and communities they consider, for whatever reason, unsafe. Theoretically, an intrusive event shows the difference between a communicative-advanced stage at which netizens have rich and verifiable information about the attack and are able to spread it widely through their social network and a communicative-backward stage at which netizens have little verifiable information about the potential risks, and can pass along the little they know mostly by word of mouth to their immediate circle of relatives and neighbors (Figure 1). They are distrustful of the government's commitment to their welfare. However, complaining about safety in one's immediate social circle is one thing, doing something about it politically in the public sphere is quite another. With overt political actions impractical and even dangerous, netizens resort to movement: they distance themselves from places of danger. Compared with sending posts, voting with their feet more convincingly shows the deep distrust of netizens about the safety claims of autocrats.

Political netizens in my study period also show deep distrust of the autocrats' safety claims (Model 2 in Table 3 and 4). Some, however, supported the claims originally, at least in the time immediately after the attack. Those who are critical online may have their views censored, a byproduct of expressing their distrust publicly. Notwithstanding these different shades of trust and distrust, I estimate that all netizens, political ones included, have moved significantly away from Kunming Railway Station and the Dashuying Uighur Community.

5.3 Hypotheses 3 Rejected and 4 Supported: After the deadly event, rich netizens are more prone to trust the information given by autocrats

Netizens are not a homogeneous entity. Which groups show distrust more? I focus on two: the powerful and the wealthy. In theory, government officials in Kunming share power with autocrats and are obliged to follow their lead in any serious crisis. As for the rich, they may also enjoy political power by virtue of connection with the autocrats. This being the case, both government officials and the rich should show more trust towards autocrats.

A difference-in-differences estimation indicates that netizens from gated communities, who are well-to-do and enjoy considerable prestige, move significantly closer to Kunming Railway Station and the Dashuying Uighur Community after the attack than do netizens of less economic means, as shown in Model 1 in Tables 5 and 6 respectively. Such a finding supports Hypothesis 3: netizens who live in gated communities and, more generally, the urbane and the rich put more trust in the information that autocrats provide in the event of an intrusive disaster than do those less economically favored. I suggest that the rich and the well-connected believe more in the willingness and the ability of autocrats to keep these places safe. Surprisingly, netizens connected with the government do not differ significantly from netizens not so connected in their distrust of officialdom to cope, as shown in Model 2 in Tables 5 and 6. Moreover, unlike rich netizens, who are trustful, netizens with links to the government show rather less trust in the information autocrats provide and believe, furthermore, that autocrat policies, by their harshness and rigidity, can lead to future terrorist incursions. Such a finding implies that government officials themselves are among those who take a jaundiced view of government information.

[Tables 5 and 6 about here]

6. Conclusion

My findings reveal the mechanisms of information contestation between autocrats and netizens in a digital era of internet, social media and smartphones. In normal, peaceful times, autocrats hold an advantage over netizens. When a sudden and threatening event occurs, they may again hold an advantage and enjoy, at least for a time, widespread popular trust and support. The response of netizens is more ambiguous and complex. Not knowing whether autocrats' information is to be fully trusted, their behavior, verbal and in action, can become erratic. On the one hand, they may see certain weaknesses in the autocrats' position that they can exploit by extracting from them data of a certain kind.

What are such weaknesses in autocrats? One is their failure to predict the intrusive event. While such failure is to be expected, it does dent the autocrats' image of omniscience and, with it, their promise of security and stability. Another challenge in the case of the Kunming Railway attack is labelling. How are autocrats to label the perpetrators? They cannot be called Uighur terrorists, for that would tarnish and offend an entire population. So they are called Xinjiang separatists, but that label is still unsatisfactory, for separatists do not necessarily kill. The label "Xinjiang separatist terrorists" may be accurate, but it is unwieldy. Then there is the challenge over the control of information, which is of primary importance to all authoritarian regimes. How is total control over information possible when the local people, including their commercial press, are almost certain to get to the scene first and what they see and report is quickly diffused by social media on smartphones. Autocrats are then left with other political tools in their kit, which is to co-opt or embellish the news that has already spread.

This research also has broader methodological and conceptual implications. Using social media to study sudden events in an authoritarian regime complements the standard approaches of

survey and interview, which can only be conducted days and even weeks after the event and are, in this sense, retrospective and historical, their data—the traces of memory—colored by the time that has lapsed. This is a problem in all social-science survey-interview research. One that is distinctive to authoritarian states is not just the retrospective bias, for one that is more serious by far is the reluctance on the part of the surveyed and interviewed to be honest in their replies, even if their memory doesn't play tricks. The scientist returns to his campus to analyze the results with the tolling bells of academia in the background. The people he interviewed who said what they really thought might forever dread the midnight knock on the door.

In addition, this research provides a new angle on social media studies. Conventionally, such studies focus on the 'media' aspect of the social media—how certain content is manipulated through propaganda or censorship. My research, by contrast, brings the 'social' aspect to the forefront—netizens versus autocrats, rather than topical contents, is the cynosure. Because ordinary netizens are not as active as journalists on social media or those who are politically engaged (Chan and Fu, 2013), this 'social' aspect tends to be neglected.

Conceptually, in a broad sense, my study contributes to the literature on authoritarian politics. Authoritarian regimes are challenged by neither formal institutions from within, such as the law courts and election, which are notoriously window-dressing, nor, in the case of China, effectively by informal institutions such as nonprofit organizations, private charities, universities, and non-hierarchical churches. Surprisingly, civil society itself, empowered by the new communication technologies and enriched by a capitalist economy that the regime itself has helped to found, do exert a pressure on the autocrats, with the result that they occasionally compromise with the rising tide of educated netizens and tiptoe toward social, though not yet political, liberalization.

Autocrats, even given the challenges listed above, have their own ways to respond. They can, as I have just said, embellish and co-opt the news the netizens possess. They can also disclose what information they have and add to it facts they know to be untrue but which they can remove later. This doling out of “facts” gives them an aura of sincerity and competency, of doing all that they can to alleviate a dire situation. But to the extent that autocrats have become addicted to manipulating information, netizens for all their surface compliance show a distrust that is more pervasive and deeper than meets the eye. Wise autocrats will take notice, but without democracy, what effective steps can they take other than by pretending to be liberal-democrats, when necessary?

REFERENCES

Ai Weiwei. 2011. *Ai Weiwei's Blog: Writings, Interviews and Digital Rants, 2006–2009*. Cambridge, Mass.: MIT Press. Edited and translated by Lee Ambrozy.

Brady, A., 2008. *Marketing dictatorship: propaganda and thought work in contemporary China.* Lanham, Md: Rowman and Littlefield.

Brady, A., 2009. Mass persuasion as a means of legitimation and China's popular authoritarianism. *American Behavioral Scientist, 53*: 434−457.

Cairns, C., 11 August 2015. *The "social media shock."* Unpublished manuscript. Cornell University Press.

Cairns, C. and Carlson, A., 2016. Real-world Islands in a Social Media Sea: Nationalism and Censorship on Weibo during the 2012 Diaoyu/Senkaku Crisis. *The China Quarterly*, 225, pp.23-49.

Chang, C., and Manion, M. 2016. Why do autocrats permit the media to play to the gallery? Information management in sina news coverage of china's anticorruption campaign. Working Paper.

Chen, J., Pan, J. and Xu, Y., 2015. Sources of authoritarian responsiveness: A field experiment in china. American Journal of Political Science.

CNNIC. 2015. *The 36th Statistical Report on Internet Development in China (Zhong-guo Hulian Wangluo Fazhan Zhuangkuang Tongji Baogao).* Beijing: China Internet Network Information Center.

CNNIC. 2016. *The 37th Statistical Report on Internet Development in China (Zhong-guo Hulian Wangluo Fazhan Zhuangkuang Tongji Baogao).* Beijing: China Internet Network Information Center.

Dunning, Thad. 2008. *Crude Democracy: Natural Resource Wealth and Political Regimes*. New York: Cambridge University Press.

Earle, J.S. and Gehlbach, S., 2015. The Productivity Consequences of Political Turnover: Firm-Level Evidence from Ukraine's Orange Revolution. *American Journal of Political Science*, 59(3), pp.708-723.

Egorov, G., Guriev, S. and Sonin, K., 2009. Why resource-poor dictators allow freer media: A theory and evidence from panel data. *American political science Review*, 103(04), pp.645-668.

Enikolopov, R., Petrova, M. and Sonin, K., 2016. Social media and corruption. Working Paper.

Freedom House. 2015. *2015 Freedom of the Press Data*, https://freedomhouse.org/report-types/freedom-press

Fu, C., Samet, H. and Sankaranarayanan, J., 2014, November. WeiboStand: capturing Chinese breaking news using Weibo tweets. *In Proceedings of the 7th ACM SIGSPATIAL International Workshop on Location-Based Social Networks* (pp. 41-48). ACM.

Fu, King-wa, Chung-hong Chan, and Michael Chau. 2013. "Assessing Censorship on Microblogs in China: Discriminatory Keyword Analysis and the Real-Name Registration Policy." *IEEE Internet Computing*, vol. 17, no. 3: 42–50.

Gandhi, J., & Przeworski, A. 2007. Authoritarian Institutions and the survival of autocrats. *Comparative Political Studies*, *40*(11), 1279–1301.

Geddes, B. 1999. What Do We Know About Democratization After Twenty Years? *Annual Review of Political Science*, *2*(1), 115–144. http://doi.org/10.1146/annurev.polisci.2.1.115

Goffman, E., 1974. Frame analysis: An essay on the organization of experience. Harvard University Press.

Greitens, S.C., 2013. Authoritarianism Online: What Can We Learn From Internet Data In Nondemocracies? *PS: Political Science & Politics*, 46(02), pp.262-270.

Gunitsky, S., 2015. Corrupting The Cyber-Commons: Social Media As A Tool Of Autocratic Stability. *Perspectives on Politics*, 13(01), pp.42-54.

Hassid, J., 2008. Controlling the Chinese media: an uncertain business. *Asian Survey*, *48*: 414–430.

Hong, J. 2011. From the world's largest propaganda machine to a multipurposed global news agency. *Political Communication, 28*: 377–393.

Kelly, S., Cook, S.G. and Truong, M., 2012. *Freedom on the net 2012: a global assessment of internet and digital media.* Freedom House.

Kennedy, J. J., 2009. Maintaining popular support for the Chinese communist party: The influence of education and the state-controlled media. *Political Studies, 57*: 517–536.

King, G., Pan, J. and Roberts, M.E., 2013. How censorship in China allows government criticism but silences collective expression. *American Political Science Review*, 107(02), pp.326-343.

King, G., Pan, J. and Roberts, M.E., 2016. How the Chinese Government Fabricates Social Media Posts for Strategic Distraction, not Engaged Argument. Copy at http://j. mp/1Txxiz1 Download Citation BibTex Tagged XML Download Paper, 2.

Knockel, J., Crete-Nishihata, M., Ng, J.Q., Senft, A. and Crandall, J.R., 2015. Every Rose Has Its Thorn: Censorship and Surveillance on Social Video Platforms in China. *In 5th USENIX Workshop on Free and Open Communications on the Internet (FOCI 15).*

Kuran, T., 1997. *Private truths, public lies: The social consequences of preference falsification*. Harvard University Press.

Landry, P.F. and Stockmann, D., 2009. Crisis management in an authoritarian regime: Media effects during the Sichuan Earthquake. *Available at SSRN 1463796.*

Lorentzen, P., 2014. China's strategic censorship. *American Journal of Political Science*, 58(2), pp.402-414.

MacKinnon, R., 2011. China's" networked authoritarianism". *Journal of Democracy*, 22(2), pp.32-46.

Magaloni, B. and Kricheli, R., 2010. Political order and one-party rule. *Annual Review of Political Science, 13, pp.123-143.*

Manion, M., 2016. *Information for Autocrats: Representation in Chinese Local Congresses.* Cambridge University Press.

Meng, T., Pan, J. and Yang, P., 2014. Conditional Receptivity to Citizen Participation Evidence From a Survey Experiment in China. *Comparative Political Studies, p.0010414014556212.*

Miller, Blake Andrew Phillip. 2016. "Automatic Detection of Comment Propaganda in Chinese Media." *Available at SSRN 2738325.*

Morozov, E., 2012. *The net delusion: The dark side of Internet freedom*. Public Affairs.

Ng, J. Q. 2014. Tracing the Path of a Censored Weibo Post and Compiling Keywords that Trigger Automatic Review. The Citizen Lab, Toronto.

Ng, J.Q., 2015. Politics, Rumors, and Ambiguity: Tracking Censorship on WeChat's Public Accounts Platform. The Citizen Lab, Toronto.

Nip, J.Y. and Fu, K.W., 2016. Challenging Official Propaganda? Public Opinion Leaders on Sina Weibo. *The China Quarterly*, 225, pp.122-144.

Noesselt, Nele. 2014. "Microblogs and the Adaptation of the Chinese Party-State's Governance Strategy." *Governance,* vol. 27, no. 3: 449–468.

O'Brien, K.J., Li, L., McAdam, D., Tarrow, S.G. and Tilly, C., 2006. Rightful resistance in rural China. Cambridge University Press.

Qin, B., Strömberg, D. and Wu, Y., 2016. The Political Economy of Social Media in China. Working paper.

The State Council of People's Republic of China, 2016, the thirteenth five-year plan for economy and social progress. http://www.gov.cn/xinwen/2016-03/17/content_5054992.htm.

Stockmann, D., 2010. Who believes propaganda? Media effects during the anti-Japanese protests in Beijing. *The China Quarterly*, 202, pp.269-289.

Stockmann, D., 2013. *Media commercialization and authoritarian rule in China*. Cambridge University Press.

Stockmann, D., and Gallagher, M., 2011. Remote Control: How The Media Sustains Authoritarian Rule In China. *Comparative Political Studies*, *44*: 436−467.

Svolik, M.W., 2012. *The politics of authoritarian rule*. Cambridge University Press.

Tang, L., and Sampson, H., 2012. The Interaction Between Mass Media And The Internet In Non-Democratic States: The Case Of China. *Media, Culture, and Society*, *34*: 457–471.

Tang, W., and Iyengar, S., 2011. The Emerging Media System In China: Implications For Regime Change. *Political Communication*, *28*: 263–267.

Yang, X., Yang, Q., and Wilson, C., 2015. Penny for Your Thoughts: Searching for the 50 Cent Party on Sina Weibo. In Ninth International AAAI Conference on Web and Social Media.

Zandbergen, Paul A., and Sean J. Barbeau. 2011. "Positional Accuracy of Assisted GPS Data from High-Sensitivity GPS-enabled Mobile Phones." *Journal of Navigation*, vol. 64, no. 3: 381–399.

Zhu, T., Phipps, D., Pridgen, A., Crandall, J. R., & Wallach, D. S. 2013. The Velocity Of Censorship: High-Fidelity Detection Of Microblog Post Deletions. arXiv preprint arXiv:1303.0597.

Table 1. Number of comments on Sina News and posts on Sina Weibo both before and after the attack

	Number of stories	Number of accessible comments	Number of Kunming smartphone netizens found commenting on stories
Five-month before	12,969	49,322	80
Five-day after	134	20,224	295

Table 2. Political, non-political netizens and the transformation between them 5 months before and 5-days amid the Kunming Railway Station Attack

		Before the Kunming Railway Station Attack		
		Political	Non-political	Total
Amid the Kunming Railway Station Attack	Political	416	1,996	2,412
	Non-political	1,124	10,819	11,943
	Total	1,540	12,815	NA

McNemar's chi-squared = 243.15, p-value < 2.2e-16

Table 3. Netizen movements around the Kunming Railway Station before and after the attack

Netizen distance to the Kunming Railway Station, km	**Model 1**	**Model 2**
Post-Attack	0.65***	0.68***
(binary)	(0.113)	(0.113)
Intercept	15.67***	15.80***
	(0.080)	(0.081)
User fixed effects	Yes	Yes
Political user controls	No	Yes

Observations: 57,081 Sina Weibo users in Kunming before and after the attack

Standard errors in parentheses, variable descriptions in text

*** p<0.001

Table 4. Netizen movements around the Dashuying Uighur community before and after the attack

Netizen distance to the Dashuying Uighur community, km	**Model 1**	**Model 2**
Post-Attack	0.68***	0.70***
(binary)	(0.115)	(0.115)
Intercept	16.13***	16.19***
	(0.082)	(0.082)
User fixed effects	Yes	Yes
Political user controls	No	Yes

Observations: 57,081 Sina Weibo users in Kunming before and after the attack

Standard errors in parentheses, variable descriptions in text

*** p<0.001

Table 5. Netizen movements around the Kunming Railway Station differentiated by connections to Government or gated community

Netizen distance to the Kunming Railway Station, km	**Model 1**	**Model 2**
Post-Attack (binary)	1.03***	0.72***
	(0.122)	(0.116)
Post-Attack × Gated community	-1.170***	
	(0.284)	
Gated community (binary)	-10.07***	
	(0.202)	
Post-Attack × Government		-0.65
		(0.418)
Government (binary)		-12.26***
		(0.289)
Intercept	17.49***	16.63***
	(0.086)	(0.082)

Table 6. Netizen movements around the Dashuying Uighur Community differentiated by connections to Government or gated community

Netizen distance to the Kunming Railway Station, km	**Model 1**	**Model 2**
Post-Attack (binary)	1.05***	0.73***
	(0.122)	(0.118)
Post-Attack × Gated community	-1.229***	
	(0.287)	
Gated community (binary)	-10.50***	
	(0.204)	
Post-Attack × Government		-0.68
		(0.424)
Government (binary)		-12.30***
		(0.293)
Intercept	17.96***	17.02***
	(0.088)	(0.083)

Figure 1. The theoretical framework of information contestation between autocrats and netizens

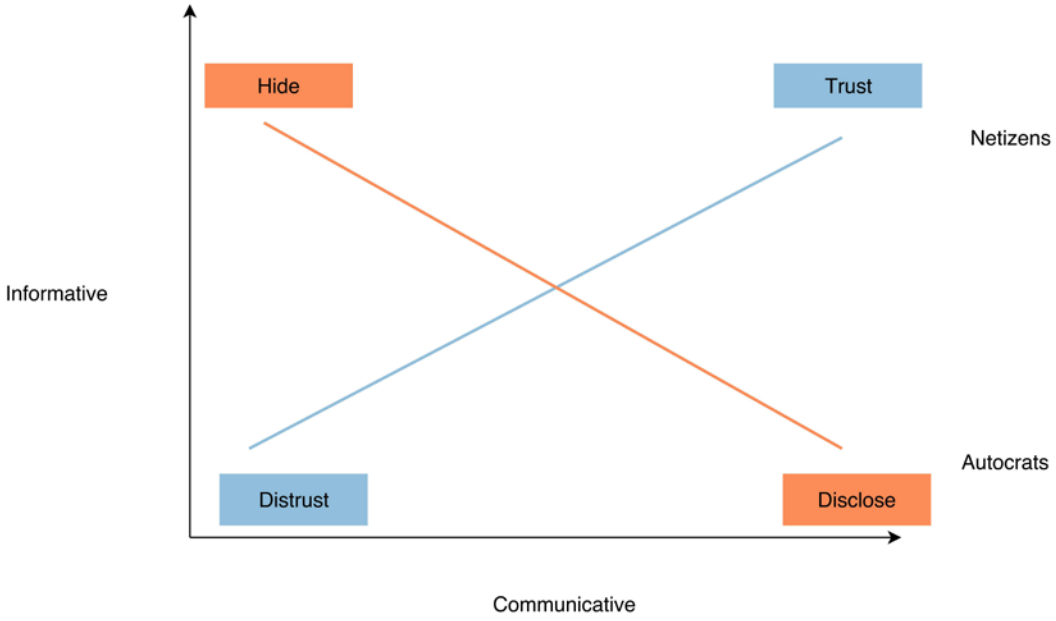


Figure 2. Number of political posts sent by hour over the event

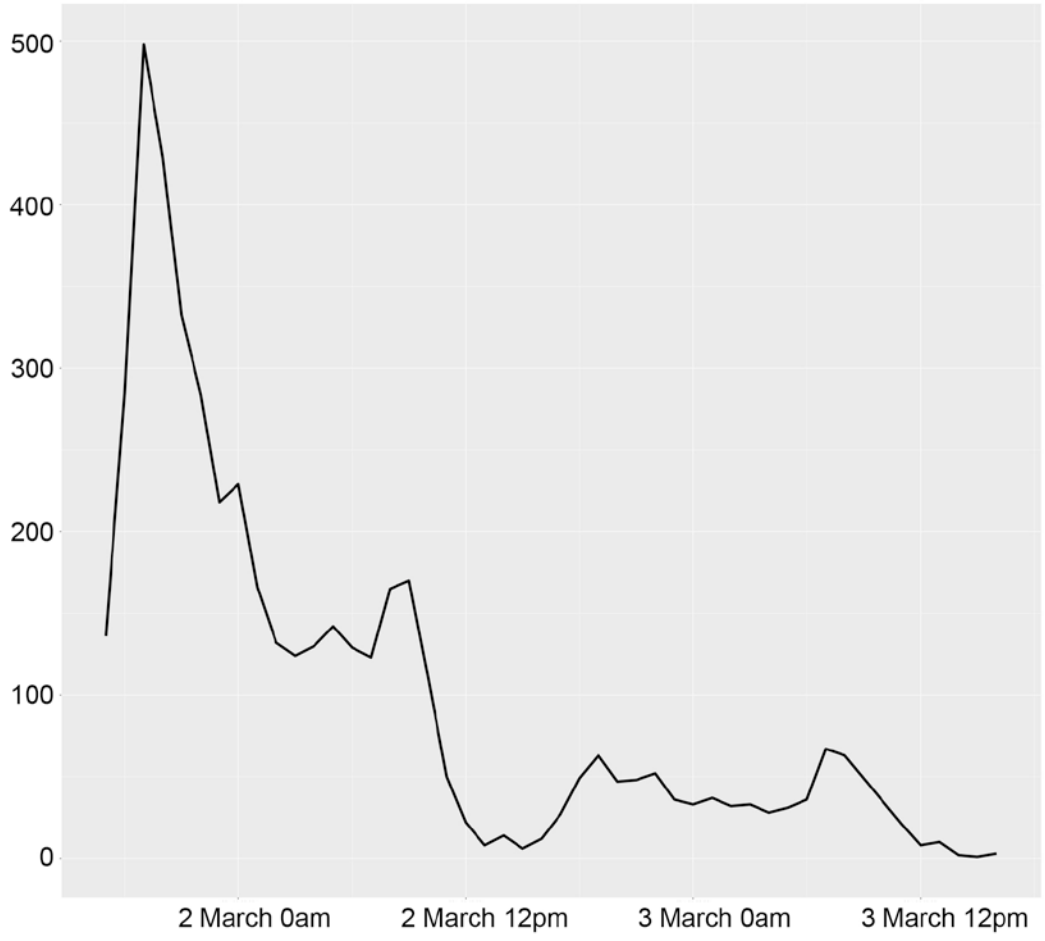


Figure 3. The avoidance of network distance illustrated near Kunming Railway Station and Dashuying Uighur Community

The shaded black line indicates the average network distance that each netizen moves away from these focal places, reflected by their Weibo posts.

Appendix A. Data Collection

I collect and store geotagged posts from netizens in Kunming on Sina Weibo, for the 10-month period from 9pm on 1 November 2013 through 9pm on 1 August 2014. This amounts to more than 1.4 million unique posts. Geotagged posts are those sent on a device, typically a smartphone, with a tag of location that the netizen choose to show with the post, as well as to the underlying precise location of the device at the time of posting. The location is highly precise: the positional error is 2 to 25 meters (Zandbergen and Barbeau 2011). I am confident that we are collecting all geotagged posts in Kunming within 10 days after the post was sent. The Sina Weibo Nearby Application Programing Interface (API) permits us, using a programming interface with customized parameters, to access the Sina Weibo database that stores all such posts. To systematically retrieve posts in Kunming, I designed a net of 10,001 location points set 400 meters apart from one another; this entirely encompasses the boundaries of Kunming municipality.[47] The Sina Weibo Nearby API limits us to retrieving from a location point, at any one time, a maximum of the 20,000 most recent posts.[48] We navigate our net of points, point by point, to retrieve the most recent posts within a radius we set at 1,000 meters. It takes us approximately three weeks to navigate every point in our net, retrieving posts at each point. By design, the radius of 1,000 meters for each of 10,001 location points set 400 meters apart from one another implies considerable spatial overlap, as illustrated below in Figure A.1. This

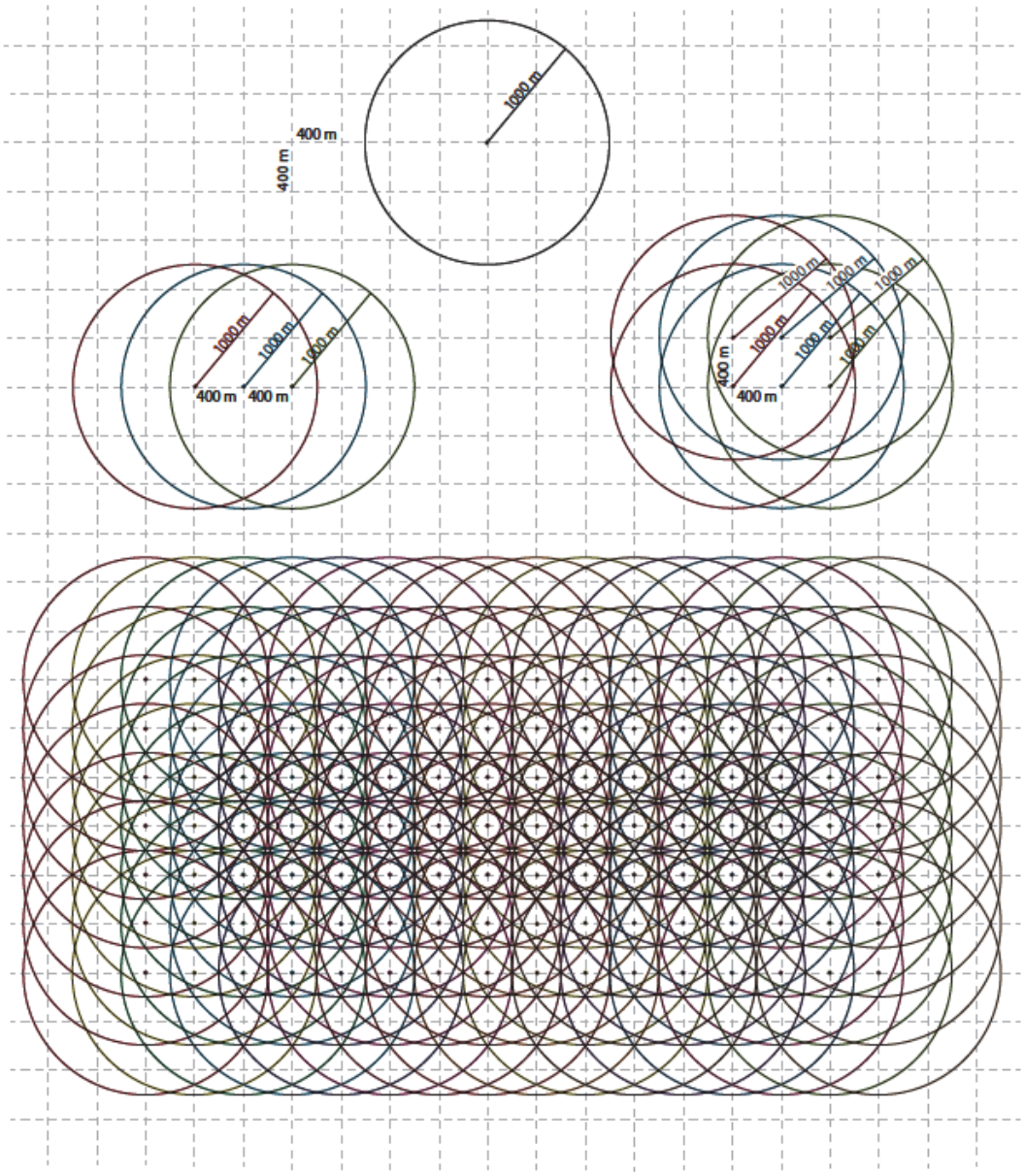
[47] It encompasses the area from 39.463 to 40.580 degrees north in latitude and from 115.815 to 117.176 degrees east in longitude.

[48] Specifically, Sina Weibo Nearby API has a technical limit of 50 posts per page and limits retrieval to the most recent 400 pages. If we view all 400 pages from a single location point at the same time, which we do here, we can view a maximum of the most recent 20,000 posts.

contributes to intended redundancy in our retrieval of posts: in each three-week period, as we navigate our net of points systematically, from point to nearby point, at each point retrieving up to 20,000 most recent posts within the 1,000-meter radius, our retrieval covers every piece of Beijing's geographic space many times over.[49] Given spatial and temporal overlap, with a net of 10,001 points located 400 meters apart and retrieval of all posts within a radius of 1,000 meters of each point every three weeks, we miss a post only if there are more than 20,000 posts from some single location in three weeks—which is effectively impossible.[50]

[49] We subsequently eliminate multiple observations of the same post to create a dataset of unique posts.
[50] Empirically, we find this limit of 20,000 posts in the retrieval space of a single location point is reached in about six months.

Figure A.1. A Net of Points in a Sea of Posts



Appendix B. Differences between Geotagged and Non-Geotagged Posts

Geotagged posts may be different from other posts, even intentionally different. If differences are significant, it threatens the validity of our research. Given my research question, I am mainly concerned about bias in the political content of geotagged posts. Are geotagged posts systematically more apolitical? Are geotagged posts with political content systematically less critical? I, along with my advisor, assess these potential threats to validity by making use of items in the BAS 2015 survey on use of geotagging and all questions about politics available to us. Specifically, we compare differences in political attitudes of netizens who report disabling the smartphone geotagging function when posting on Sina Weibo and netizens who report never having done so.

Among the 2,610 Beijing residents that the BAS 2015 surveyed, 1,609 (62 percent) are netizens. Thirty percent of them go online to access Sina Weibo—most of them at least several times a week, many of them several times a day. Smartphones are by far the most popular technology: 94 percent of BAS 2015 microbloggers go online by smartphone. We ask all netizens who have ever posted anything on Sina Weibo by smartphone if they know that smartphone posts can be geotagged: 92 percent of them say yes. Moreover, 39 percent of these geotag-aware microbloggers report they have disabled geotagging for at least some posts. How (if at all) do these microbloggers differ politically from those who report they have never disabled geotagging when posting on Sina Weibo? It seems plausible that those who have disabled geotagging are relatively more politicized, more politically sophisticated, perhaps more cynical about politicians and their official pronouncements.

To investigate this issue, we analyze differences between the two groups in terms of their interest in politics and views on the anticorruption campaign, as reflected in responses to five

survey questions. Response frequencies are shown in Table B.1. We find no statistically significant differences between those who report having turned off geotagging and those who report never having done so. Based on their responses, the two groups do not differ significantly in: their interest in politics, how much the anticorruption campaign affects their lives, their agreement with rhetoric about the campaign's purpose and the sincerity of efforts to fight corruption, or their opinion on whether or not the campaign should continue and intensify.[51]

[Table B.1 about here]

We also conduct a thorough search for other differences between the two groups. We find statistically significant differences on only two dimensions. First, microbloggers who have disabled geotagging are more highly educated: 64 percent of them have attended university, compared to the 53 percent of microbloggers who have never disabled geotagging.[52] Second, microbloggers who have disabled geotagging have more online followers: a reported 620 followers on average, compared to 280 for microbloggers who have never disabled geotagging.[53]

These results suggest that our focus on geotagged posts does not pose a threat to validity for our research. The two groups do not differ on the dimension of interest to us: political views. Moreover, the two differences of significance suggest that the netizens whose posts we analyze are more "ordinary" than very highly educated or more popular netizens. To the extent that this presents a bias, we find it acceptable. We purposely aim to shift attention from the political

[51] Specifically, the Pearson chi-square probability for observed differences in the distribution of responses across the two groups is higher than .05 (and usually much higher) for all five items. With this many observations, the Pearson chi-square test is very powerful, that is, highly likely to point to differences as statistically significant.

[52] This is not about the difference between university students and others, however. The difference in proportions currently studying, rather than working, is trivial and not statistically significant.

[53] For highest level of education, a categorical variable in the survey, we use a Pearson chi-square test; for number of online followers, we use a t-test of difference in group means. For both tests, we adopt .05 as our level of statistical significance.

activists and celebrities that are the subject of most studies of the social media in authoritarian settings to focus instead on more ordinary netizens.

Why have 39 percent of microbloggers at some time disabled geotagging when posting on Sina Weibo? We offer respondents a choice of eight reasons (and the opportunity to give another reason) for disabling geotagging. By far the most common choice, making up 39 percent of responses, is: “I don’t want strangers to know my location.” The second most common choice, accounting for 14 percent of responses, is: “I don’t want friends, family, or co-workers to know my location.” In short, in their online (inherently public) activity, Beijing netizens sometimes seek an element of privacy—not from Sina or the state, which is impossible in any case, but from other netizens.

Table B.1. Beijing Microbloggers Who Have (or Have Never) Disabled Geotagging, Percentages

	Geotag on	Geotag off
Interest in politics		
Very interested	3.9	5.4
Somewhat interested	28.9	27.0
Not too interested	59.5	62.2
Not interested at all	7.8	5.4
Impact of current campaign on your life		
Direct impact	11.5	11.0
Indirect impact	4.1	5.5
Basically no impact	84.4	83.5
Current campaign: highly important or high-level political struggle?		
Highly important	56.5	53.2
High-level political struggle	17.7	14.9
Both	19.1	29.1
Neither	6.7	2.8
Fighting corruption: response to mass public or political show?		
Reflects public sentiment, enjoys public support	64.7	60.3
Political show	7.0	5.9
Both	21.9	30.9
Neither	6.5	2.9
Approve of continuation and intensification of campaign?		
Very much approve	57.5	63.3
Somewhat approve	32.9	27.2
So-so	7.9	9.5
Somewhat disapprove	1.8	0
Very much disapprove	0	0

Notes: Percentages compare netizens who report having disabled the geotagging function for some posts when posting on Sina Weibo with netizens who report never having done so. Frequencies are responses to the following questions. (1) Are you interested in politics? Would you say you are very interested, somewhat interested, not too interested, or not interested at all? (2) Do you think the current anticorruption effort has an impact on your own life? A direct impact? An indirect impact? (3) Some view the current anticorruption effort as highly important:

not to counter corruption may destroy the party and country. Others view it as a high-level political struggle. Which view do you agree with? (4) Some think fighting corruption reflects the public sentiment and enjoys broad popular support. Others think fighting corruption is a political show. Which view do you agree with? (5) Do you approve of the government continuing and intensifying the anticorruption effort?

Source: Beijing Area Study 2015

Appendix C. Information manipulation on media and social media, including Sina News and Weibo

Authoritarian regimes have been known as information manipulators on cyberspace (Morozov 2012, Egorov et al. 2009, Kelly et al., 2012, Greitens 2013, Lorentzen, 2014, Noesselt 2014, Gunitsky 2015, Enikolopov et al, 2016), but no regimes can compete with the single-party China in the scale and extent of such manipulation. The Chinese Communist Party (CCP) presides over one of the most formidable information management operations in the world, a program "unprecedented in recorded world history" in size and sophistication (King, Pan, and Roberts 2013, 326). Freedom House ranked China 186th of 199 countries concerning press freedom (Freedom House, 2015), where the effective use of the media as a political tool is always a top priority (Tang and Iyengar 2011). Scholars have reach a consensus that—through their propaganda bureaucracy, Chinese autocrats do in fact effectively manage news reporting to bolster authoritarian rule (Brady 2008, 2009, Hassid 2008, Kennedy 2009, Hong 2011, Stockmann and Gallagher 2011, Tang and Sampson 2012, Stockmann 2013).

Chinese autocrats also extend their media control to social media. As the Great Firewall of China blocks the most prominent social media sites, such as Facebook and Twitter, Chinese netizens are largely restricted to the Chinese social media, outstandingly the public microblogging platform Sina Weibo and the private messaging platform Tencent WeChat[54]. A growing body of literature has studied the party's control of Chinese social media,

[54] Sina Weibo occupies 70% of microblogging market according to CNNIC report (CNNIC 2015), such that Weibo stands for Sina Weibo in China. In contrast, Tencent WeChat are mostly private, equivalent to Tencent QQ as a messaging or group messaging service.

overwhelmingly focusing on censorship. Apparently, as Chinese social media companies that targets Chinese media consumers, every company is not unscathed to the variety of known censorship: from petite to severe are search blocking, credit-point penalty taking, post publicity limiting (i.e. prohibition for reposting and/or commenting, or limit viewership to private circle), post automatic reviewing, post keyword filtering, post manual deletion, account suspension, account deletion, public humiliation, and personal jailing (MacKinnon 2011, Knockel et al. 2015, Ng 2014, 2015).

Known evidences suggest that netizens have enjoyed considerable freedom of speech on social media, in comparison to the stern control of journalists on media. However, autocrats still control the resources—a large number of social media account with millions of followers, and are able to set the framing of political discourse (Nip and Fu, 2016). Studies on censorship find that censorship mostly applies on social media celebrities, journalists, and political activists with a wide variation of censorship rate, not ordinary netizens, for political posts from celebrities have a much greater publicity. Focusing on 280 thousand Weibo users, Chen found that up to 50% of political posts from Weibo celebrities were censored, compared to overall 1% censoring rate including other politically active netizens (Chen et al. 2013). Focusing on another 3,567 political activists and celebrities on Weibo, Zhu et al. found that the censoring rate was at staggering 12% (Zhu et al. 2013), which is close to the censoring rate of 13% found by King, Pan, Roberts[55]. In comparison, following around 300 thousand users with more than 1,000 followers, WeiboScope project at the University of Hong Kong found the censoring rate was at a negligible level of 0.01% (Fu, Chan, and Chau 2013). Another study of 13.2 billion posts from a marketing

[55] Study by King, Pan, Roberts mostly focuses on a variety of BBS, the social media with much longer history.

company also finds millions of posts on protests and corruptions that are uncensored (Qin et al 2016). My survey results on 2,620 ordinary netizens confirm this negligible censoring rate for smartphone Weibo users who have the habit of posting with geotag—most netizens have never seen any post deleted.

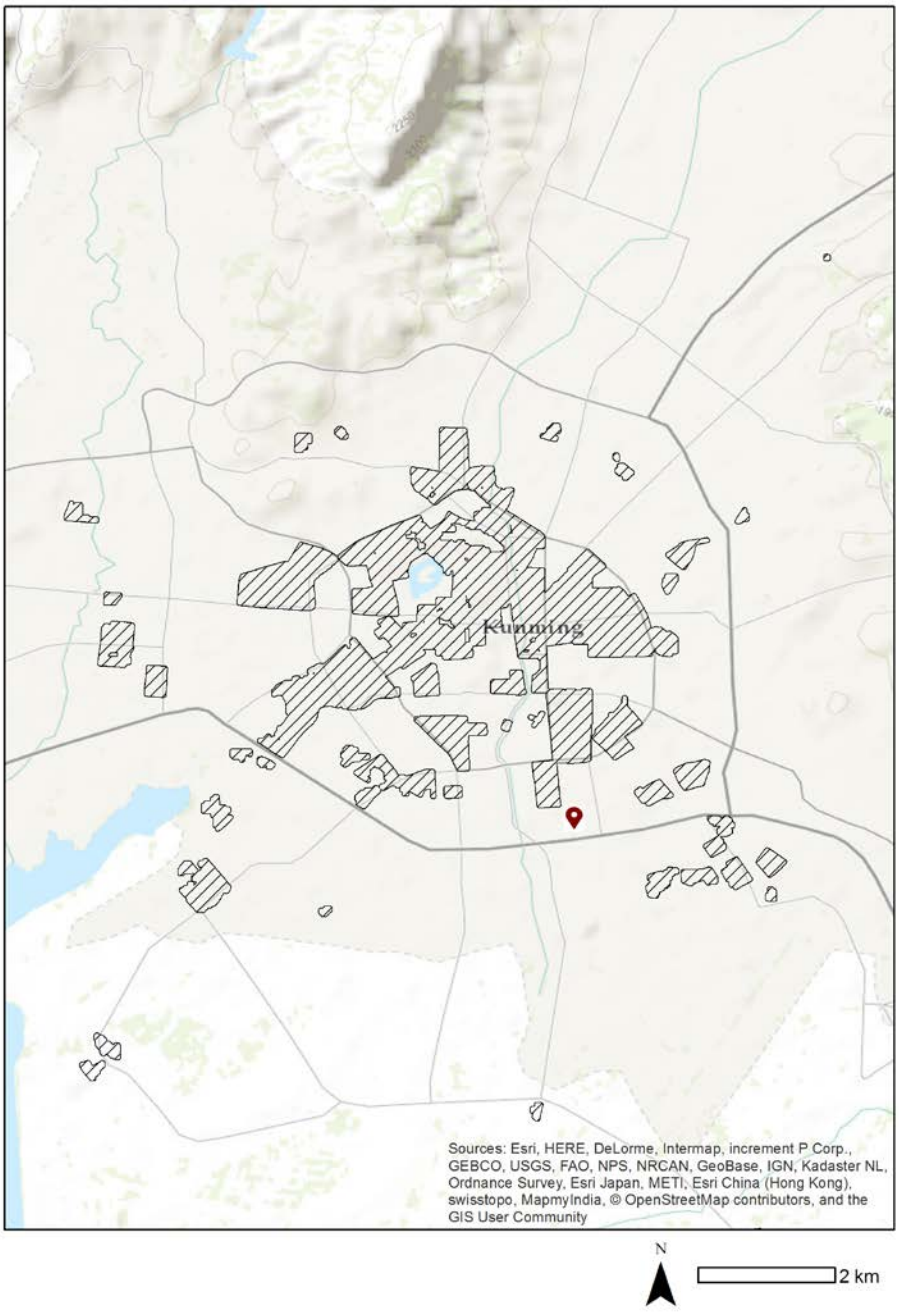
In addition to censorship, autocrats fabricate netizen engagement of their political framing by employing paid-commentators. King, Pan, Roberts (2016) estimate that paid-commentators have posted 488 million social media posts every year on various social media sites. Miller (2016) find paid-commentators are active in guiding public opinion over sudden intrusive event, such as Tianjing Explosion. However, some studies also find strong netizen engagement from authentic users over political events. Strong engagement has been seen over the sovereign dispute of the Diaoyu/Senkaku Islands on Weibo (Cairns and Carlson 2012). And in a study to find paid commentators on Weibo, Yang et al found no clusters of paid-commentators; instead they found clusters of ordinary netizens supportive to the CCP regime (Yang et al. 2015).

Appendix D. The Space of government and gated community

I use the POIs (points of interest) from Soso Map to identify the government office and residence from my result of part one. From the 4,445 POIs that related to government, I only choose the nearly a thousand POIs that represent Yunnan provincial bureaus (e.g. Yunnan Public Security Bureau) and Kunming municipal bureaus (e.g. Kunming Inland Revenue Bureau). These bureaus are provincial and municipal levels often enjoy guarded spaces exclusive to the officials and their affiliates. The offices and residences that have geographical intersection with any of the selected POIs are labeled as governmental entities. To further eliminate geometric gaps from my result of part one, I buffer the governmental entities 100-meter by each geographical feature, shown as the shaded area in Figure D.1. I then use the shaded area to identify netizens with government connections.

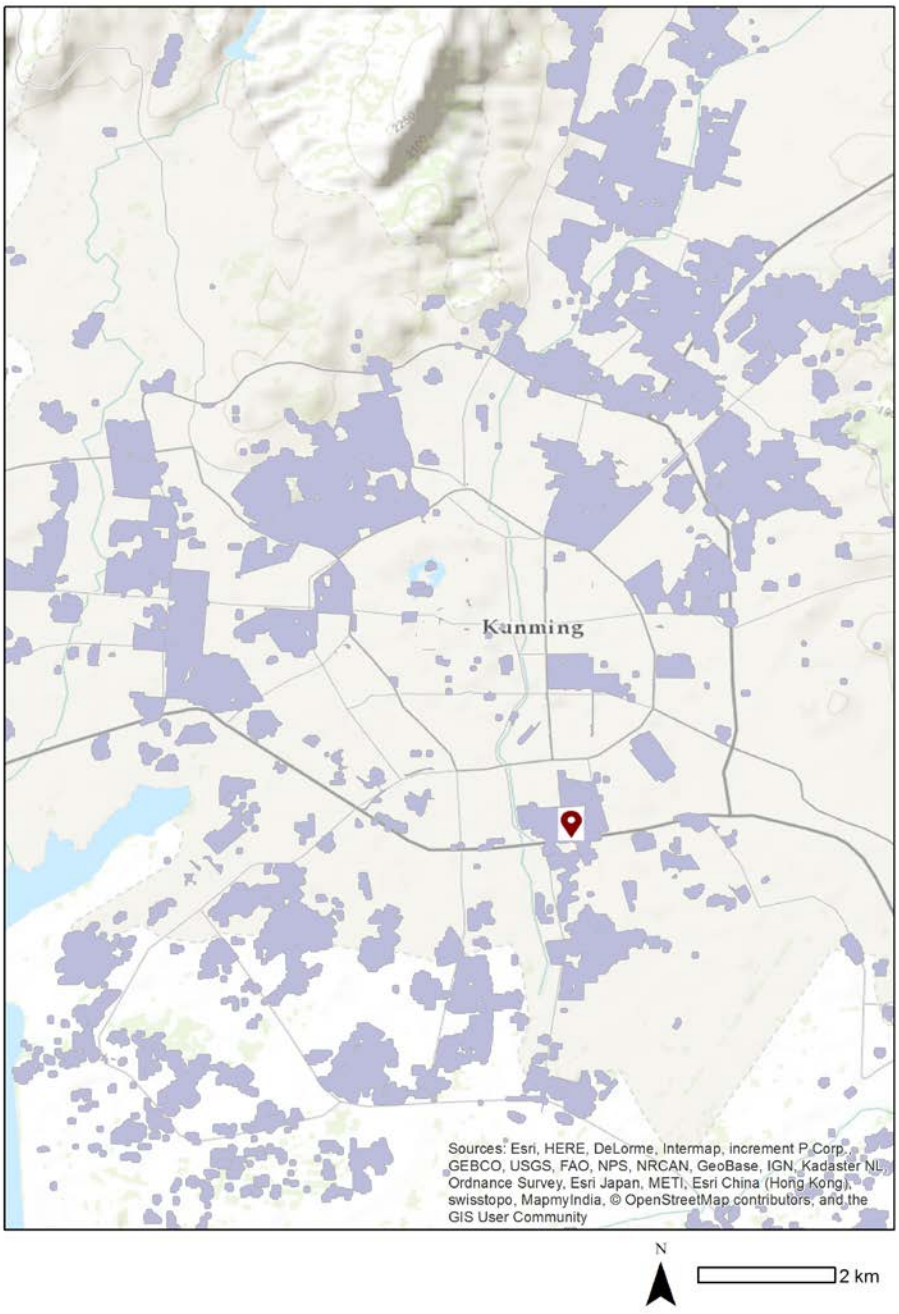
I also define the gated community consistent with my result of part one. I remove the gated communities that are also government residences, and then buffer 100-meter to close geometric gaps. I use the buffered area, shown as the purple areas in Figure D.2, to identify netizens as those in gated community.

Figure D.1 The governmental space in Kunming



The shaded areas are the spaces of government discussed in the paper. The red place mark indicates the Kunming Railway Station where the attack occurs.

Figure D.2 The gated communities in Kunming



The purple areas are the gated communities. They are in different architectural types, such as

skyscrapers, building-complex, single-family range house, or a combination of apartment buildings. The red place mark indicates the Kunming Railway Station where the attack occurs.