2. Social unrest in China: a bird's-eye view

*Christian Göbel**

IMAGING SOCIAL UNREST IN CHINA

Judged by protest frequency estimates, China is a country where protests have become routine. In 2011, a study published by Landesa survey claimed that, according to "Chinese researchers," in 2010 China saw "180,000 mass incidents . . . 65 percent of them related to land disputes."[1] Drawing on this figure, a headline in *The Atlantic* made the misleading claim that "500 protests [occurred] every day." Despite the fact that little is known about how the unnamed "Chinese researchers" cited in the Landesa survey arrived at this figure, most publications, including that of the author, refer to it.[2] However, these figures contribute little to a better understanding of social unrest in China. Instead, they conjure up the powerful image of a China in serious turmoil. If we take the results of a Google Images search as a representation of how the public mind might imagine "protests" or "social unrest" in China, we see hundreds of angry people holding up banners, raising their fists, and shouting slogans. Some photos show combat-ready riot police, and more extreme images feature overturned cars and the use of teargas canisters.

While such scenes certainly occur, it is unclear whether or not they are representative of most instances of social unrest in China. Our knowledge about protests in China is sketchy: the media tends to cover large-scale events, which, as will be shown, are very infrequent; and academic research is mainly based on small-N case studies of protesters[3] and of local officials.[4] Fewer attempts have been made to understand how protests influence governance in China at large;[5] and only a handful of studies provide insights into the spatial and temporal distribution of protests in China, which grievances they address, how many people they involve, and how likely protests are to meet with repression.[6]

A macro-perspective of social unrest in China is needed to judge the impact of social unrest on the stability of China's one-party authoritarian regime. For example, while a violent protest that draws a large crowd is a major challenge to the regime, a gathering joined by only a handful of people is not. Low attendance signals a lack of public interest and assures the authorities that a grievance can be safely ignored.[7] Also, demands for financial compensation are relatively easy to defuse by, for example, "buying stability," whereas civil rights protesters cannot be bought off so easily.[8] Finally, protests can be an indicator of specific local problems; but they can also highlight systemic deficiencies, for example when grievances are not confined to a particular locality or region.[9]

Drawing on a dataset of 74,452 protests that occurred in China between 2 June 2013 and 13 June 2016, this chapter provides a bird's-eye view of social unrest in China by addressing the issues just outlined. The results show that most protests are nothing like the image evoked above would suggest: protests in China are widespread but tend to occur seasonally and involve fewer than 30 participants. Most protests are recorded in the days before Chinese New Year, when factories close their accounts and migrant workers return home. Financial compensation, not substantive rights, are at the heart of

Christian Göbel - 9781786433787
Downloaded from Elgar Online at 10/28/2019 07:38:30PM
via free access

most protests. Another noteworthy finding is that the number of protests against land grabs and evictions has remained stable at a low level, while those staged by (prospective) homeowners against real estate developers and property management companies have increased steeply.

Homeowners represent, as Jean-Louis Rocca puts it, "the politically conscious vanguard of the Chinese middle classes," which means that a steep rise in such protests has important implications for China's political and social stability.[10] As David Goodman points out, however, 88 percent of households in urban China own at least the leasehold to their own home, so homeownership does not necessarily equate to middle-class status.[11] Still, their activism closely matches the behavior the literature on political transformation associates with the middle classes. According to Rocca, homeowners "disseminate laws and regulations protecting their rights, as well as newspaper articles criticizing developers and management companies. In their eyes, the problem comes from the local authorities whose power is still beyond control."[12] Rocca points out that, as of present, homeowners are not interested in political change and, at least in Beijing, are "incorporated in the wheels of bureaucracy"; but widespread discontent among this group will have to be taken seriously.[13]

The findings discussed so far might suggest that the authorities have only the increase in homeowner protests to worry about. According to the logic espoused above, most of the small protests can be ignored, and the participants in larger ones can be compensated. However, the findings highlight some additional trends that should concern the leadership in Beijing. For one, protesters frequently take their grievances to the capital, which is one of the hotspots where protests occur. Second, the fact that labor-related protests are both frequent and widespread suggests that there exists a major systemic problem that has not yet been resolved and might endanger stability in the long run. Third, it is well known that protesters engage in "troublemaking" because their demands will be ignored otherwise. However, by allowing themselves to be goaded into cracking down on such protests, the authorities demonstrate to onlookers how fundamentally authoritarian China is. Consequently, an event that would otherwise not have concerned bystanders is turned into a display of how local governments, who struggle to cast themselves as caring, responsive and service-oriented, routinely act against the legitimate interests of China's citizens.

THE DATASET

As much as we know about social unrest in China from the perspective of protesters and local officials in individual locations, we still lack a macro-perspective of the phenomenon. Representative statistical data that could shed light on this issue does exist, but suffers from important limitations. Such data comes in two forms: (a) survey data indicating whether survey participants have ever witnessed or participated in a protest; and (b) protest event data collected by the China Labor Bulletin[14] and the political scientist Manfred Elfstrom.[15] All of these datasets are detailed and of high quality, and they cover several years. However, they record only a few thousand cases, which greatly limits analysis at the subnational level. Furthermore, the protest event data is confined to labor issues and, more specifically, strikes.

Christian Göbel - 9781786433787
Downloaded from Elgar Online at 10/28/2019 07:38:30PM
via free access

To gain a more comprehensive insight into social unrest in China, this chapter draws on an archive of protest-related "tweets" collected and published online by Lu Yuyu, who is also known by his alias "Wickedonna." Each entry is a collection of different sources, most prominent among which are eyewitness accounts in the form of Weibo tweets by eyewitnesses. Several pictures are attached to each entry, thus increasing the credibility of the source. The posts span the period from July 2013 to June 2016, when Lu and his girlfriend Li Tingyu were arrested on charges of "picking quarrels." More than one year later, on 3 August 2017, Lu was sentenced to four years in prison,[16] while Li Tingyu had been released from captivity in June that year without an official court verdict having been issued.[17]

Lu, who sustained himself by means of donations,[18] had made collecting protest posts a full-time occupation. He manually searched Weibo and other social media platforms for relevant material, examined each post for its veracity, and published the results on his website.[19] All entries were downloaded, and meta-information such as the date and location of a protest and the estimated number of participants was extracted with the help of dictionaries and regular expressions. The resultant database contains tweets documenting more than 70,000 protests, roughly 29,000 of which took place in 2015. The textual data was then coded and analyzed with the help of unsupervised and supervised machine learning algorithms, which will be briefly introduced in the relevant sections.

The database entries provide information on the reason for a protest and sometimes also on the type of collective action engaged in by participants, for example demonstration, strike, or riot. In some cases posts are composed not by participants but by observers. Most observers seem to sympathize with the protesters, but some also complain about the inconvenience that a protest caused them. Finally, around one-third of all entries provide clues on how the local authorities reacted to a protest. The posts do not usually reveal whether or not the demands of the protesters were met, but often mention the presence of (armed) police, public security, urban law enforcement officials (*chengguan*) or hooligans allegedly hired by local officials to intimidate protesters: 16 percent of all entries mention that protesters were beaten by security forces or hired hands.

Since the protest accounts have been hand-collected by one person, it is possible that the dataset is not representative of all protests in China. Three factors in particular could cause selection bias: censorship, the predispositions by Lu Yuyu, and a geographical bias caused by different degrees of Internet penetration across China.

Regarding censorship, H. Christoph Steinhardt convincingly illustrates that protests are now more intensely discussed by Party leaders in public. Moreover, censorship of protest reports in the media has significantly loosened up, not least as a consequence of "more assertive Internet and news media."[20] As for Lu's choice of protests for inclusion on the website, comparing the entries to existing scholarly works, it appears that the database is quite representative of the protests described in the print media and on the Internet. Finally, since most entries constitute social media information supplied by eyewitnesses, it is possible that protests occurring in underdeveloped regions, where Internet penetration is low, were less likely to make it onto the webpage. The presence of a geographical and technological selection bias was ruled out by regressing the log number of protests in a city against a dummy variable for each province and the city-level broadband penetration rate. When controlling for a number of structural variables—gross regional product (GRP) per capita, contribution of the primary sector and foreign direct investment

Christian Göbel - 9781786433787
Downloaded from Elgar Online at 10/28/2019 07:38:30PM
via free access

(FDI) to GRP, government expenditure per capita, population density, average wage level, banking deposits per capita, enrolment in pension, and medical and unemployment insurances—broadband penetration and the provincial dummies lose their significance.

A BIRD'S-EYE VIEW OF PROTESTS IN CHINA

Drawing on the data just described, the following sections provide insights into the grievances that cause social unrest, the spatial distribution of protests in China, their size, seasonal variation, and likelihood of crackdown.

Protest Issues

This section investigates the grievances that motivated the protests recorded in the database, revealing many surprising findings. First of all, the proportion of land-related protests, which the Landesa survey estimated to constitute the absolute majority of all protests, is comparatively small and follows a downward trend. Instead, protests by homeowners are on the rise. Also surprising is that environmental protests, which have gained significant media attention in recent years, are few and far between. The largest category by far is labor protests, which account for more than 40 percent of all events in the dataset. Finally, the data reveals that very few protests are related to the protection of civil or political rights, and the database contains not a single anti-regime protest.

To gain some initial insights into the structure and contents of the corpus, I used Mallet, a Java-based natural language processing software,[21] to create "topic models," i.e. weighted lists of words that frequently co-occur in a corpus (topics) and scores indicating the relative weight of each topic in a document (document-topic scores).[22] Each topic takes a value between 0 and 1, and all topics add up to 1. Short documents describing only a grievance and not much else will score highly on only one topic, whereas long entries that contain more information will score intermediate values on several topics. Each topic in turn consists of a number of terms more or less related to that topic, with each term also receiving a score to indicate its weight.

The creation of meaningful topic models is not a simple matter; the researcher needs to decide on a number of parameters, such as: the choice of algorithm; words that should be excluded; how often the corpus should be iterated over; the threshold for considering a model as "converged"; whether some topics should be allowed to be stronger than others; and, last but not least, the number of topics that should be produced. All this is a matter of trial and error and assumes familiarity with the corpus. For the present analysis, I generated a model with 30 topics (T1–30) by letting the algorithm iterate 10,000 times over the corpus with an optimization step every ten iterations.

Table 2.1 shows the results of the analysis. It lists the ten most frequent terms for each topic along with an average document topic score which indicates the weight of a topic in the corpus. Although these scores are useful for estimating which topics figure more prominently than others, they should not be misinterpreted as a measure of the frequency of a topic. To recap, if a labor protest is mentioned in a short tweet, the topic score for "labor" will be higher than for a longer tweet that provides additional information. While either describes a labor protest, the topic score of the former will be higher than that of

Christian Göbel - 9781786433787
Downloaded from Elgar Online at 10/28/2019 07:38:30PM
via free access

Table 2.1 Topic model of Wickedonna corpus

ID	Topic strength	10 most frequent terms
1	0.16	migrant workers wages workers hard-earned money pay talks arrears boss migrant workers home government
2	0.15	government people petition people people leadership rights call the shots district government video
3	0.14	traffic jam avenue trouble transportation highway police road vehicle traffic jams banner
4	0.13	police hit SWAT assault video ordinary people government people police station violence
5	0.10	forwarding government hope people help friends video attention thankyou fair
6	0.09	people government hope society national rights leadership interest maintenance support
7	0.09	owners developers rights submitted house real estate cell problem commitment deception
8	0.08	banner rights square Wanda protest video broke news trouble users official
9	0.05	villagers farmers land government land forced forced recruitment people wounded mafia
10	0.05	owners district property rights developers residents garden poly heating problem
11	0.05	live events issues represent department representative understand staff processing leader
12	0.04	employees strike workers company boss limited wages factory rights compensation
13	0.04	demolitions demolition villagers government house houses demolition placement church cross
14	0.03	villagers land secretary corrupt officials corruption rights government village village leaders village
15	0.03	family members dead dead bodies children killed police station hospital family police
16	0.03	government protest marches police SWAT people out demonstrations live masses
17	0.03	government regulations construction national problem resettlement compensation projects immigration policy
18	0.03	company investment hard-earned money the government funding limited guarantee fraud bank victims
19	0.02	merchants rights rent protest shops collective shopping business owners rent
20	0.02	taxi strike drivers companies taxi collective rental bus black car bus
21	0.02	workers workers company employees group business leadership wages rights limited
22	0.02	rights Wanke pan-Asian mobile petitioner national site protest Tanglong video
23	0.02	urban management beat law enforcement people police crowd violence live events police
24	0.02	children parents school kindergarten elementary school school student department of education education school district
25	0.02	hospital family members doctors people children death patients patients surgery rescue
26	0.02	pollution residents garbage villagers government protest chemical plants waste incineration projects people
27	0.02	villagers government masses reservoir housing coal mining farmers leaders people
28	0.01	teacher teacher student school wages strike rights college highschool collective
29	0.01	court attendance membership use report comments lawyer applications attention register
30	0.01	veterans people government SWAT rights war railway people Zhongtai people

Christian Göbel - 9781786433787
Downloaded from Elgar Online at 10/28/2019 07:38:30PM
via free access

the latter. For example, labor-related topics (T1, T21, T28) together account for around 20 percent of all document-topic scores; but, as will be seen later in the chapter, more than 40 percent of all protests are related issues.

The topics fall into three groups. The first highlights the reasons for protests and will be discussed below. The second (nine topics) contains various activities and descriptions of states of affairs, including reports that petitioners have assembled in front of the district or city government (T2), complaints about protest-related traffic jams (T3), and terms describing large-scale protests (T16). The first two topics, which perhaps identify the tweeters as onlookers rather than participants, account for 29 percent of the corpus. Tweets that urge Internet users to help a cause by retweeting constitute another 10 percent (T5). This topic nicely illustrates how people are employing social media to further their causes. Phrases urging the government to protect the interests of the people (T6) are equally common, followed by T8, another topic that probably fits into the "onlooker" category. The remaining topics are relevant in only a fraction of the corpus. They include words related to the legal process (T29), phrases (T11), labor-related synonyms (T21), and a residual category (T30).

The third group consists of two law enforcement-related topics. Both include the names of law enforcement agencies as well as vocabulary related to acts of violence. T4, which combines the noun "police" with verbs such as "hit," "assault" and "violence," indicates that tweeters mainly associate the police with acts of repression (and not, say, protection). That this topic accounts for 13 percent of the total topic score suggests that police brutality occurs not infrequently. T23 refers to both urban management officials and the police, and also contains violence-related terms. As will be discussed in more detail, other topics contain such terms as well.

The following subsections introduce the grievance-related topics in more detail. It has already been established that labor issues are the main trigger for social unrest in China. They fall into four broad groups, among which protests by migrant workers, an especially vulnerable group in China's labor market, are the most common (T1).[23] Other topics simply mention "workers" (T12, T21) and refer to the plight of taxi drivers (T20) and teachers (28). It is well known that taxi drivers go on strike because they are dissatisfied with the government-regulated fare system, which does not compensate drivers for sudden increases in the price of gasoline. As oil prices plummeted in 2014, this is less an issue in the dataset than the unfair competition by illegal taxis ("black cabs"). In recent years, taxi strikes have been aimed at forcing local leaders to fight illegal taxis and protect cabbies from the competition by ride-sharing companies such as Didi Chuxing. School teachers are another status group that occasionally engages in social unrest, the main reason being wage arrears. Because of local budget shortages, teachers are frequently owed months of salary by their public employers, are not paid as much as they are entitled to, or fail to receive the subsidies they had been promised.

Property-related protests constitute the second largest category overall and can be subdivided into three distinct groups: land grabs and forced evictions; conflicts related to the sale of real estate; and dissatisfaction with services provided by property management companies. Different from what might have been expected from both the Landesa survey and the existing literature, protests related to land grabs and evictions are not prominent in the dataset, but homeowner protests are.[24] While it is possible that previous research has overestimated the significance of land grabs and evictions as drivers of social unrest,

Christian Göbel - 9781786433787
Downloaded from Elgar Online at 10/28/2019 07:38:30PM
via free access

it is not likely. It seems more reasonable to assume that such protests have indeed declined, which might well be the result of disincentives set by the central government. Citizens are now better protected against wanton land grabs and evictions than in the early 2000s, and compensation is fairer than it was.

Still, T9, T14, T13 and T17 illustrate that there is still much room for improvement. As becomes clear from T9, land reacquisition takes the form of forced expropriation so frequently that the algorithm fails to separate the nature of the grievance from the heavy-handed means applied by local authorities. As can be seen, words like "forced," "forced appropriation," "wounded" and "mafia" figure prominently in T9. That officials in rural China and companies hire ex-convicts or organized gang members to intimidate or use force against people who resist land grabs or development projects is well established in the literature, so the presence of terms related to hooliganism makes sense.[25] T14 also refers to land grabs, but puts more stress on allegations of corruption. Sometimes protesters seek to obtain justice not by directly voicing an instrumental demand, but by accusing the village authorities of corruption.[26] T27 might also be related to this subcategory of property-related issues.

Eviction-related terms dominate T13 and T17, each of which highlights a different aspect. T13 is more about the owners, whereas in T17 terms related to construction projects and the question of resettlement are more prominent. Tweets loading high on T13 tend to describe protests against an impending eviction, while those falling into T17 come from people who have already abandoned their homes but are meeting with difficulties during resettlement. Examples here are relocations from the city center to the outskirts, and apartments that are too small or too expensive. It is interesting to note that the decapitation or complete demolition of church buildings in Zhejiang province also feature in T13.

It seems that, in terms of property-related protests, homeowners might have taken the place of the victims of land grabs and evictions. T7, which describes conflicts between real estate developers and buyers, constitutes the second largest grievance. Judging from the tweets that load on this topic, the most prevalent issue seems to be that buyers fail to receive their ownership certificate despite having already made a down payment on a new apartment. Frequently, the construction of an apartment complex has fallen behind schedule, and owners complain that they cannot move in. Other tweets report how developers have gone bankrupt and buyers have lost all their investments. In some cases, protesters occupy the construction site, often still a vacant space, and demand that construction continue. Dissatisfaction with developers cutting corners or erecting buildings where a garden had initially been planned also falls into this category.

The third category of property-related grievances is related to the availability, quality and cost of utilities (T10). Property management companies are often accused of negligence—examples in this case being elevators in disrepair, low water pressure or deficient heating systems. Charges for services that should be free also motivate homeowners to stage demonstrations.[27] While much ink has been spilled on land grabs and evictions, the far more frequent conflicts between buyers and developers are comparatively under-researched and constitute a promising avenue for future research.

The topic related to environmental protests (T26) is surprisingly weak given how prominent environmental issues are both in the news and in China-related research.[28] Remarkably, people hardly ever protest against air and water pollution, which is a

Christian Göbel - 9781786433787
Downloaded from Elgar Online at 10/28/2019 07:38:30PM
via free access

major grievance in most cities. In fact, most environmental protests are of the "not in my backyard" (NIMBY) kind, with protesters mobilizing against the construction of waste incinerators or chemical factories near their homes.[29] That said, the paucity of environmental protests should not be interpreted to mean that environmental issues do not matter to Chinese people: they tend to draw large and heterogeneous protest crowds, and are therefore more likely to turn into anti-regime protests if the situation escalates.[30]

T24 consists of terms related to education. Education-related protests are initiated by parents, students and teachers. Parents mostly protest against the impending closure of a kindergarten or a primary school, or, in the case of migrants, that their children are excluded from the education system altogether. High school and university students tend to take to the streets because they perceive university entrance examinations as being unfair to them. Finally, as mentioned above, wage arrears are a common reason for protests by teachers. Surprisingly, only very few protests are initiated because of well-documented problems such as teachers charging irregular fees or forcing students to take costly private classes,[31] a phenomenon still commonplace, especially in rural China.[32]

Social unrest because of medical mistreatment (T25) and the death of a family member (T15) is the subject of two further topics. These topics correlate to some extent because family members sometimes assemble in front of a hospital to vent their anger over the death of a loved one. The term "treated to death" (*zhisi*) frequently appears in tweets in this category. However, there are also events that are unique to each category. With respect to medical mistreatment, fake medicine or high bills are common reasons for people taking to the streets. The tweets belonging to the death of a family member category are even more heterogeneous—people demand justice after a family member has died in a workplace or traffic accident, a student has jumped to her death, or a person has drowned in a pond. Two kinds of event load heavy on this topic: protests resulting from the death of a family member, or the death of a family member resulting from a protest. Some protests were started because urban management officials or police officers allegedly killed someone during an interrogation or while trying to contain a protest.

Finally, the analysis also identified topics related to protests initiated by investors claiming they were defrauded when purchasing financial products (T18); and by shopkeepers and hawkers complaining about sudden rent increases, not having their deposit returned, or being chased away by urban management officials (T19).

GEOGRAPHY OF PROTESTS

So where in China do protests occur? Are they confined to individual regions, or are they a nationwide phenomenon? Understanding the regional distribution of protests is important for any assessment of China's social and political stability. For example, protests that occur in only a small number of localities are less likely to link up with other protests.[33] Further, it is far easier for a government to remedy grievances and shortcomings in one location than in the whole country. In a similar vein, the geographical distribution of protests related to a particular grievance provides important clues as to whether a problem is systemic or isolated. With regard to labor protests, for example, Manfred Elfstrom and Sarosh Kuruvilla conclude that even though labor protests occur especially often in the Pearl River Delta, "a distinct regional pattern is less pronounced than previously

Christian Göbel - 9781786433787
Downloaded from Elgar Online at 10/28/2019 07:38:30PM
via free access

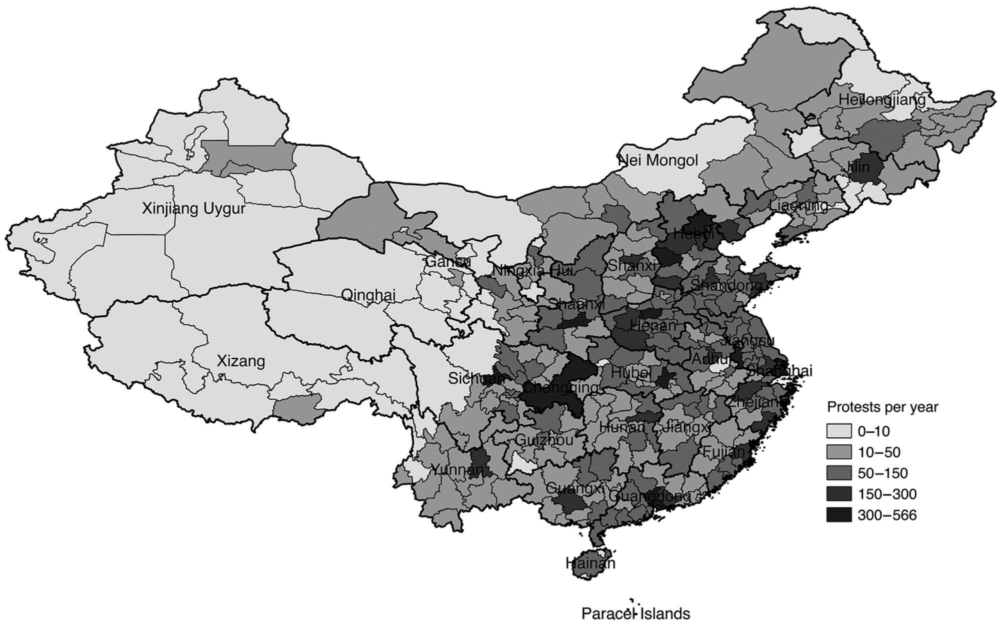

Figure 2.1 Number of protests per year

suggested."[34] In other words, labor-related grievances are not particular to manufacturing hubs in Guangdong, but are common everywhere in China. The following paragraphs will expand the analysis by accounting for other grievances besides labor.

Figure 2.1 illustrates the annual number of protests that have taken place in each prefecture-level city between July 2013 and June 2016: the darker the field, the more protests. The map confirms that protests indeed occur almost everywhere in China, although they are more common in some regions than others. This might accurately describe the reality in a location, but might also be the result of censorship. Furthermore, it should be kept in mind that the outbreak of a protest is the function of at least two factors: a grievance serious enough to motivate people to take to the streets, and the capacity of local elites to prevent this from happening.[35] It is very likely that the paucity of observed protests in Xinjiang, Tibet and Qinghai can be attributed to this last factor. Security forces have been so present in the period under observation that it would have been extremely risky and difficult to initiate social unrest.

Still, the map shows that most cities experienced, on average, at least 50 protests per year in the period of observation. As can also be seen, there are several cities that stand out for their high number of protests, most importantly Beijing, Xi'an, Shenzhen, Chongqing, Chengdu, Zhengzhou, Guangzhou and Shanghai.[36] These localities experienced between 365 and 577 protests per year. If only labor protests are included, the map (not shown here) confirms the findings by Elfstrom and Kuruvilla: labor issues are the primary reason for protests in most Chinese cities; and they are especially common in the Pearl River Delta, but also in Beijing and Chongqing. This confirms once more how serious and systematic the violation of (migrant) worker interests in China is.

Christian Göbel - 9781786433787
Downloaded from Elgar Online at 10/28/2019 07:38:30PM
via free access

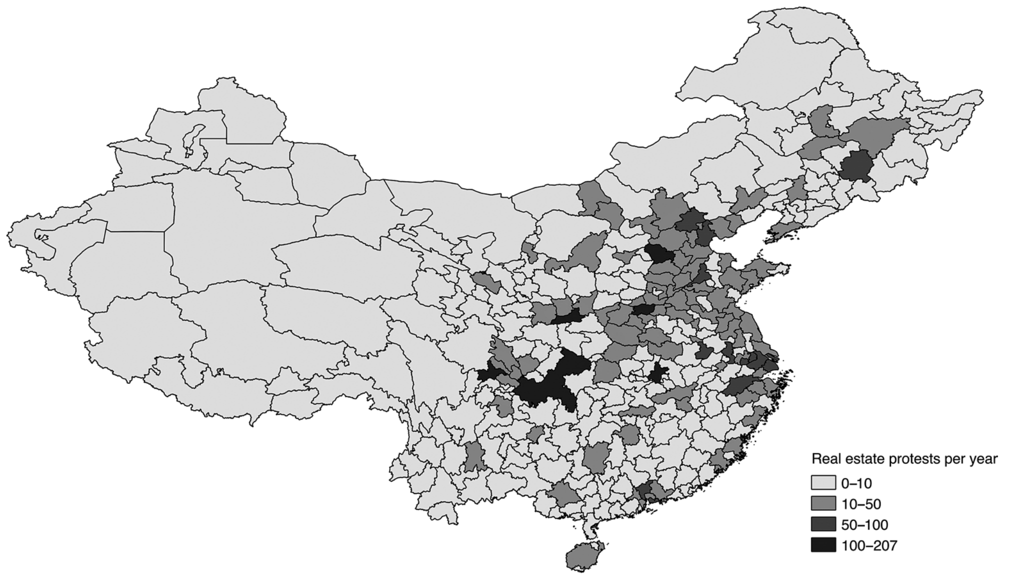

Figure 2.2 Real estate protests per year

By contrast, protests related to real estate issues are regionally more concentrated. As Figure 2.2 illustrates, they are prevalent in the coastal provinces of Zhejiang, Shanghai, Jiangsu and Shandong, but also in the inland provinces of Hebei, Henan and Sichuan. Land grabs, the third most common reason for protests, also display a regional pattern—they are mainly concentrated in southern China (Figure 2.3).

Figure 2.3 highlights a very interesting phenomenon: the cities where such protests occur often form a cluster that crosses provincial boundaries, and they share at least one border with other cities where such protests occur. The cluster spans three provinces: Guangxi, Guangdong and Fujian. As opposed to migrant worker protests, which are evenly spread across all of China, land grabs seem to constitute a more local problem particular to southern China.

CROWD SIZES

A total of 180,000 protests per year would be an imposing figure only if each of these protests were as big as we might imagine them to be based on a Google Images search. Put differently, 180,000 protests attended by 15 persons each would be less impressive than the same number of protests each attended by thousands of people. If a cause manages to rally the population of a small city, the authorities are well advised to address this cause because the same or even a larger number of people might take more extreme measures if their demands are not satisfied.[37] Conversely, there is no incentive for the authorities to take action where there are only a few protesters and no onlookers. A low turnout might signal to the authorities that an issue can be safely ignored.[38]

Christian Göbel - 9781786433787
Downloaded from Elgar Online at 10/28/2019 07:38:30PM
via free access

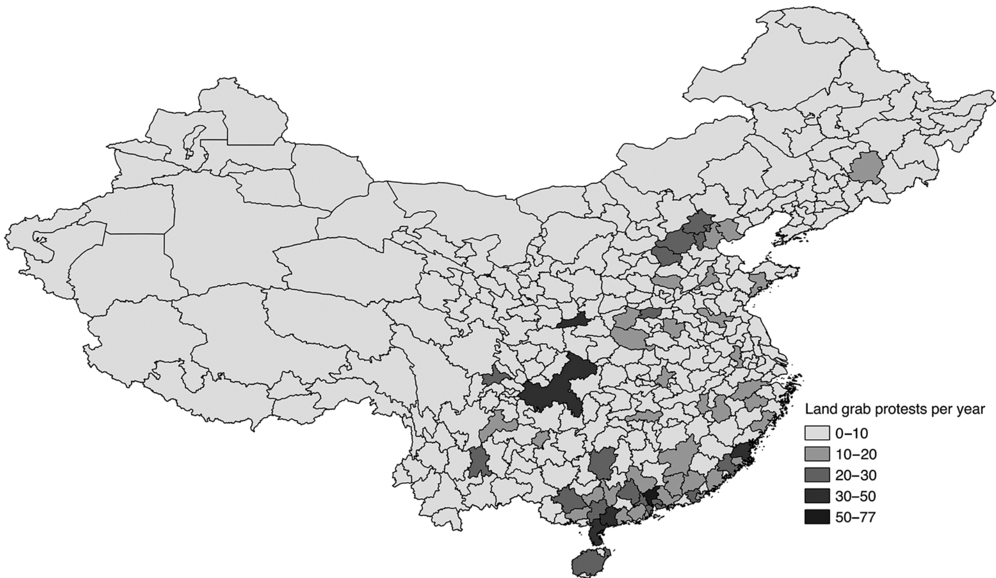

Figure 2.3 *Land-related protests per year*

That said, estimating protest sizes is notoriously difficult, the more so for large-scale protests, and especially if no aerial photos are available that would allow grid-based information.[39] Unfortunately, the only information we have are crowd estimates by eyewitnesses. A Boolean search query employing regular expressions was used to extract such information. Specifically, it searched for numbers in the immediate neighborhood of verbs denoting participation in protest activities and nouns describing different categories of participants. Unfortunately, such data is available only for 8,044 of the 74,452 events in the database. The result is displayed in Figure 2.4, which shows that there is a logarithmic relationship between the size of a protest and its likelihood to occur. In simple terms, small protests occur exponentially more frequently than large protests. The overwhelming majority of protests in China are very small, mustering fewer than 50 participants. Still, more than 2,000 events were believed by participants to have been attended by 1,000 persons or more.

Figure 2.4 also indicates a characteristic that is particular to rough estimations—they become less precise as crowd sizes increase. If only a couple of dozen people participate, estimates are often very precise because it is possible to count the number of participants. The more people there are, the more difficult it is to guess precisely how large the crowd is—for example to estimate a crowd to number 500 people rather than 400, 300 or 200. The same is true for a count of 1,000, which might in fact span an attendance of between 500 and 2,000 persons. Another spike, not visible in the graph, occurs at 10,000 persons which, according to the logic just explained, might indicate a crowd size of anywhere between 1,000 and 100,000 persons.

Overall, the overwhelming majority of protests in China average between ten and 100 participants. Large-scale protests of more than 1,000 people are exponentially less likely to occur. Yet these large protests are more likely to be reported in the Chinese and overseas

Christian Göbel - 9781786433787
Downloaded from Elgar Online at 10/28/2019 07:38:30PM
via free access

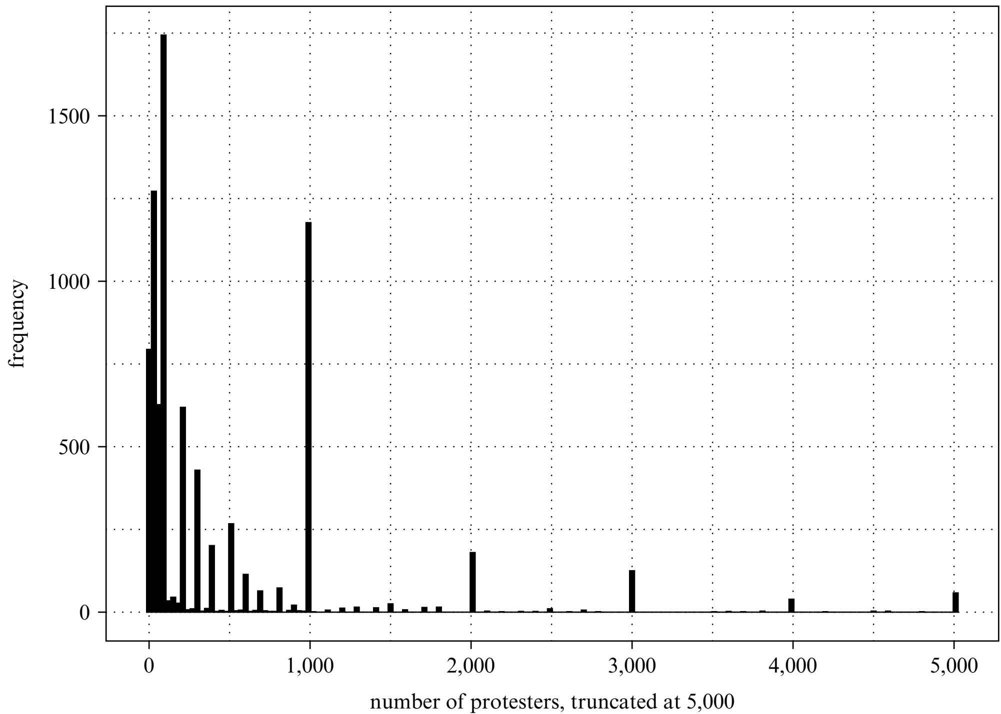

Figure 2.4 Distribution of protest sizes

media, thus forming our image of social unrest in China. In reality, the more likely a protest is to be reported in the news, the less likely it is that a protest of that size occurs frequently. Most protests in China go unreported, and perhaps even unnoticed, signaling not a weakness of the regime but that social groups are demanding to be heard.

PROTEST FREQUENCY

It makes a difference if a country experiences a steady flow of protests, as the "500 protests a day" headline in *The Atlantic* might suggest, or if protests display ebbs and flows. A sudden and unexpected burst of protest activity would be considered a national emergency, while a steady flow might signal that protests are tolerated to some degree by those in power.[40] Our data suggest that there is large variation in the number of protests that take place on any given day. The number fluctuates between 100 and 200 in spring and fall and during the winter months; it then builds to a peak of between 400 and 500 events just after the new year, then falls back into the 100–200 range. Not coincidentally, the peak coincides with Chinese New Year.

A closer look at the data suggests a likely explanation for this phenomenon: as evidenced in Figure 2.5, the overwhelming majority of protests, on average more than 40 percent, concern labor-related issues—in particular claims by migrant workers that their

Christian Göbel - 9781786433787
Downloaded from Elgar Online at 10/28/2019 07:38:30PM
via free access

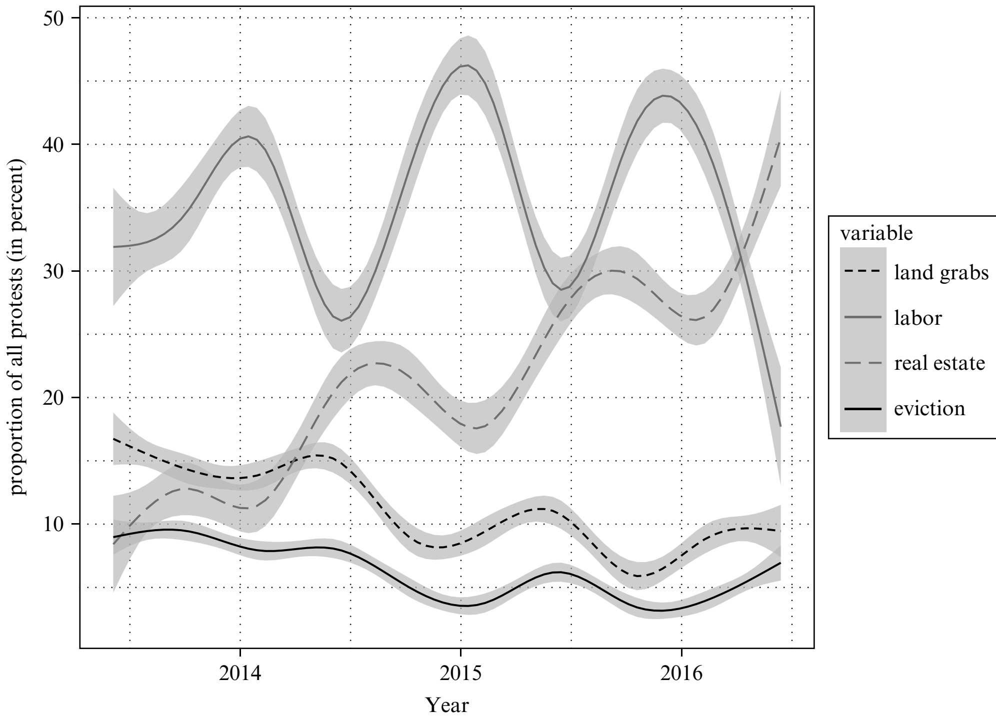

Figure 2.5 Seasonality of major grievances

wages have not been paid fully by their employers. It follows that most protests occur at a special time, i.e. the closing of accounts and the start of a new hiring season, and have a specific aim, i.e. the fair compensation of (migrant) workers for their labor.[41]

This might explain why so many protests are unable to gather followers. As important as migrant workers have been for China's economic development, they have not been able to mobilize others in support of their cause. This is not surprising—migrant workers constitute a marginalized group that members of the middle class tend to look down on, and see little reason to support. The situation might be different if labor protests were routinely repressed. China's citizens witness not only the demands made by the protesters, but also how the regime interacts with those who take to the streets. Protesters might not succeed in mobilizing China's citizens for their demands; but if their actions meet with repression, they at least help expose the violent side of local government. The repression of legitimate interests flies in the face of the image of responsive and "service-oriented" government that the CCP wishes to nurture, and might concern onlookers more than the violated interests themselves.[42] This is perhaps one of the reasons why, as will be seen in the next section, the authorities tend to act carefully when confronted with labor unrest.

Another interesting tendency is the apparent decline of land-related unrest. Land protests made up around 17 percent of all protests in 2013, but only 10 percent in 2016. A similar trend pertains to eviction-related protests, which have also declined. This does not necessarily mean, however, that land protests have become less frequent. Possibly, the decline in

Christian Göbel - 9781786433787
Downloaded from Elgar Online at 10/28/2019 07:38:30PM
via free access

land-related protests has been effected by an increase in another category of protest. Real estate-related grievances, which display a steep incline, seem to be a good candidate.

Indeed, an inspection of the absolute number of protests reveals that land- and eviction-related events have remained constant at around five incidents per day with no discernible seasonal variation. The number of real estate protests, on the other hand, has increased more than five-fold within only three years. Throughout 2013, they had been less common than land protests, but were approaching the mark of 30 protests per day in June 2016.

The steep increase of social unrest within the period of observation is perhaps the most relevant finding this study has to offer. As outlined in the previous section, such protests differ from most others in that they are driven by homeowners, many of whom are members of the middle class, and others perhaps aspiring to join their ranks. Also, many of these protests are initiated by persons who have invested a significant amount of money in their new home, so their grievance is significant. As is the case with environmental issues, such grievances pose an enormous risk to the regime because they might affect how China's various social groups relate to each other.

According to structuralist theory,[43] protests by peasants and the working class often have the effect of aligning the middle class more closely with the regime because the latter might be confronted with demands for income or property redistribution if the poor manage to seize political power. If the regime is perceived as being unable to protect the interests of the middle class, however, alliances cutting across classes become possible. Even though the specific interests of homeowners and workers might not align, they might become united in their desire for a regime that is more responsive to their interests. Along with environmental issues, real estate-related grievances pose a formidable challenge to the legitimacy of China's one-party regime.

PREVALENCE OF ISSUES

This final section investigates the prevalence of various grievances. A Support Vector Machine (SVM) was used to arrive at a more accurate classification of protest events than the one discussed earlier. SVMs are supervised machine-learning algorithms that are known to perform well in supervised document classification. In an SVM, the terms in a corpus are represented in a multi-dimensional vector space, and the association with a particular topic is assessed by calculating the distance between terms and specific vectors. A training set of documents provides the benchmark for the classification, which consists of the calculation of support vectors that delineate the spaces between clusters of vectors. New documents are classified based on their distance from these support vectors. The model was trained with a hand-coded sample of 1,600 documents, 800 of which were randomly chosen and 800 more which were added after a first run of the model. The latter set of documents was purposely chosen to fill gaps where the first round did not yield enough training material for a certain topic.[44] Eventually, I arrived at 12 different categories of protest that are largely consistent with the grievance-related topics identified by the topic model. The quality of the SVM was assessed by splitting the 1,600 documents into a training and a test set, and calculating precision and recall values for each category: 75 percent of the documents were randomly selected for training, and 25 percent for testing; and the procedure was repeated 1,000 times to arrive at robust estimates of the

Christian Göbel - 9781786433787
Downloaded from Elgar Online at 10/28/2019 07:38:30PM
via free access

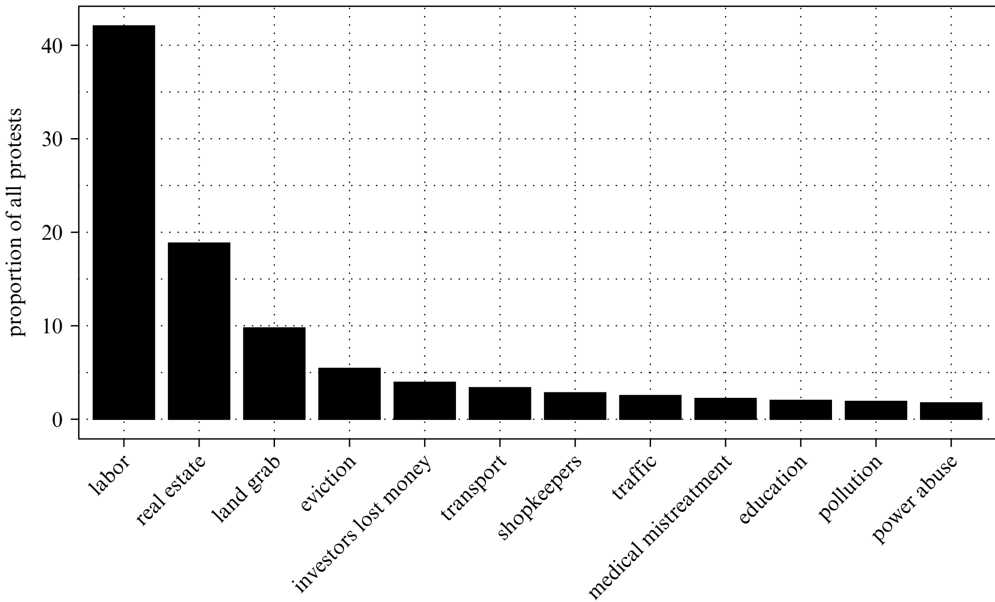

Figure 2.6 Prevalence of issues

scores. The result is very encouraging—all but two categories yield precision and recall scores above .90, and no category falls below .85. This is a very high level of convergence, especially for such a large number of categories.

Figure 2.6 illustrates the prevalence of various kinds of protest in China. As indicated earlier, labor issues are by far the most common motivator for protests, and a whopping 42 percent of all protests in the sample are related to unpaid wages, arrears and other labor issues. Real estate-related grievances motivate almost 20 percent of all protests. As explained, this category includes conflicts between owners and developers as well as between owners and property management companies. It should also be kept in mind that such protests increased rapidly over the period of observation, both in absolute and relative terms. The average displayed in Figure 2.6 betrays the fact that real estate-related grievances motivated only 10 percent of all protests in mid-2013, but 40 percent in mid-2016.

Land grabs and evictions together account for only 15 percent of all protests, which is less than expected given the prominence of the topics in the literature. Other issues motivate only a fraction of all protests. These include environmental protests; protests against power abuse; medical mistreatment; social unrest as a result of unequal access to or insufficient quality of education; and protests by defrauded investors, dissatisfied shopkeepers or bus and taxi drivers.

CONCLUSION

Drawing on a dataset derived from social media representations of popular protest, this chapter has provided a bird's-eye view of social unrest in China. It was shown that protests

Christian Göbel - 9781786433787
Downloaded from Elgar Online at 10/28/2019 07:38:30PM
via free access

are widely spread across China, but primarily occur around the Chinese New Year; and that protests are motivated mainly by instrumental demands, first and foremost the settling of wage arrears. Other grievances such as the expropriation of land and buildings or environmental pollution, are less frequent than expected, whereas protests related to real estate occupy an unexpectedly large—and increasing—part of the protest landscape in China.

These findings have important implications for our understanding of social unrest in China. First, against the background of recent scientific and journalistic concern with social unrest in China, it needs to be kept in mind that the causes of social unrest are less heterogeneous than might be assumed.[45] Labor protests are by far the largest category of social unrest, which means that investigations into the impact of social unrest on regime stability in China need to closely study China's precarious working conditions.

Second, while valuable case studies on grievances related to real estate do exist, the fact that such grievances have increased to become the second most prevalent cause of social unrest in China is both surprising and significant. The fact that members of the middle class are now habitually taking to the streets has important implications for regime stability in China.

Third, the analysis has also revealed that most of the protests are small in scale, but have a propensity to turn violent. Protesters are beaten up in around 16 percent of all protests. It follows that it might not primarily be the grievances themselves that pose a threat to stability in China, but rather the escalation of violence when protesters are beaten up by security forces or hired thugs.

Although the data used in this chapter improves our understanding of the macro-dimension of social unrest in China, it suffers from limitations. First, we do not know how representative the data really is. If the estimation of 180,000 protests in 2010 is correct, it seems likely that my dataset captures only a fraction of all protests. Future studies must establish whether this data is representative of the universe of cases or whether it suffers from selection bias. Second, the data is subjective because it represents the observations and attitudes of people who have either participated in or observed a particular event. Third, it is impossible to separate tweets by participants in a protest from those written by spectators. If, for example, the latter oppose a particular protest, the event will be misrepresented. Fourth, given the large number of small-scale protests in China, this study is biased towards such protests. Not enough data was available to study aspects like the relationship between affectedness and protest size, or protest size and protest outcome.

These issues notwithstanding, the present chapter shows the potential for analyzing event data derived from social media reports. Hopefully, more pieces will be added to the puzzle in future studies.

NOTES

* The research for this chapter was funded by the European Research Council (Grant No. 678266). I am grateful to Simon Musgrave for introducing me to word embeddings, and to David Goodman, Li Jie, Christoph Steinhardt and Andrew Wedeman for valuable comments on a previous draft. Lu Yuyu and Li Tingyu have done a great service to science by collecting the data this chapter is based on, and I acknowledge their sacrifice.

1. See Chris Buckley, "China Villagers Defy Government in Standoff over Death," Reuters World News, 15

Christian Göbel - 9781786433787
Downloaded from Elgar Online at 10/28/2019 07:38:30PM
via free access

December 2011, https://www.reuters.com/article/us-china-unrest/china-villagers-defy-government-in-standoff-over-death-idUSTRE7BE0VR20111215.

2. Christian Göbel and Lynette H. Ong, "Social Unrest in China," Long Briefing, Europe China Research and Academic Network (ECRAN) (2012).
3. See, for example, Kevin J. O'Brien and Lianjiang Li, *Rightful Resistance in Rural China* (Cambridge: Cambridge University Press, 2006); Elizabeth J. Perry, *Challenging the Mandate of Heaven: Social Protest and State Power in China* (Armonk, NY: M.E. Sharpe, 2002); Lianjiang Li and Kevin J. O'Brien, "Protest Leadership in Rural China," *The China Quarterly* 193 (2008): 1–23; Xianwen Kuang and Christian Göbel, "Sustaining Collective Action in Urbanizing China," *The China Quarterly* 216 (2013): 850–71.
4. See, for example, Peter Lorentzen, "Designing Contentious Politics in Post-1989 China," *Modern China* 43, no. 5 (2017): 459–93; Peter Lorentzen, "Regularizing Rioting: Permitting Public Protest in an Authoritarian Regime," *Quarterly Journal of Political Science* 8, no. 2 (2013): 127–58; Xiaowei Gui, "Handling of Small-Scale Protests in China: Process Dynamics and Outcomes," (Master's Thesis, Københavns Universitet, Department of Cross-Cultural and Regional Studies, 2015); Yanhua Deng and Kevin J. O'Brien, "Relational Repression in China: Using Social Ties to Demobilize Protesters," *The China Quarterly* 215 (2013): 533–52; Ching Kwan Lee and Yonghong Zhang, "The Power of Instability: Unraveling the Microfoundations of Bargained Authoritarianism in China," *American Journal of Sociology* 118, no. 6 (2013): 1475–508; Yongshun Cai, *Collective Resistance in China: Why Popular Protests Succeed or Fail* (Stanford: Stanford University Press, 2010); Yongshun Cai, "Power Structure and Regime Resilience: Contentious Politics in China," *British Journal of Political Science* 38, no. 3 (2008): 411–32.
5. For a book-length monograph on this issue, see Xi Chen, *Social Protest and Contentious Authoritarianism in China* (Cambridge: Cambridge University Press, 2012).
6. Exceptions are Elfstrom and Kuruvilla's study on labor unrest in China (Manfred Elfstrom and Sarosh Kuruvilla, "The Changing Nature of Labor Unrest in China," *ILR Review* 67, no. 2 (2014): 453–80); Ong's analysis of newspaper reports on social unrest more generally (Lynette H. Ong, "Reports of Social Unrest: Basic Characteristics, Trends and Patterns, 2003–12," in *Handbook of Research on Politics in China*, ed., David S.G. Goodman (Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, 2015), 345–60); and Chen Chih-Jou's research on the structure of protests (Chih-Jou Jay Chen, "Demanding Justice: Rising Popular Protests in China," in *Facing an Unequal World: Challenges for Global Sociology*, ed., Raquel Sosa Elízaga (New York: Sage, 2018), 250–63; Chih-Jou Jay Chen, "Growing Social Unrest and Emergent Protest Groups in China," in *Rise of China: Beijing's Strategies and Implications for the Asia-Pacific*, eds., Hsin-Huang Michael Hsiao and Cheng-Yi Lin (London and New York: Routledge, 2009): 87–105; and on protest policing in China Chih-Jou Jay Chen, "Youxiao zhili de zhigu: Dangdai Zhongguo jiti kangzheng yu gojia fanying [The shackles of effective governance: collective resistance and state reactions in contemporary China]," *Taiwan Shehuixue* 33 (2017), 113–64).
7. On the issue of credibility and rebellions, see Ravi Bhavnani and Michael Ross, "Announcement, Credibility, and Turnout in Popular Rebellions," *Journal of Conflict Resolution* (2003): 340–66.
8. Lee and Zhang, "The Power of Instability."
9. Lorentzen, "Regularizing Rioting."
10. Jean-Louis Rocca, "Homeowners' Movements: Narratives on the Political Behaviours of the Middle Class," in *Middle Class China: Identity and Behaviour*, eds., Minglu Chen and David S.G. Goodman (Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, 2013), 110–34, at 111; see also Jean-Louis Rocca, *The Making of the Chinese Middle Class: Small Comfort and Great Expectations* (Berlin: Springer, 2016).
11. This point was made in email correspondence with the author.
12. Rocca, *The Making of the Chinese Middle Class*, 186.
13. Ibid., 188. For a compatible view on middle-class protests related to garbage incinerators, see Andrew Wedeman, "Not in My Backyard: Middle Class Protests in Contemporary China" in *The Middle Class in Emerging Societies: Consumers, Lifestyles and Markets*, eds., Leslie L. Marsh and Hongmei Li (London: Routledge, 2015), 200–222.
14. http://maps.clb.org.hk/strikes/en.
15. https://chinastrikes.crowdmap.com/.
16. http://www.clb.org.hk/content/lu-yuyu-and-li-tingyu-activists-who-put-non-news-news.
17. https://www.nchrd.org/2017/02/li-tingyu/.
18. http://foreignpolicy.com/2014/04/03/meet-chinas-protest-archivist/.
19. Ibid.
20. H. Christoph Steinhardt, "From Blind Spot to Media Spotlight: Propaganda Policy, Media Activism and the Emergence of Protest Events in the Chinese Public Sphere," *Asian Studies Review* 39, no. 1 (2015): 119–37; H. Christoph Steinhardt and Zhao Litao, "From 'Stability Overrides Everything' to 'Social Governance': The Evolving Approach to Social Order in China," in *China Entering the Xi Jinping Era*, eds., Zheng Yongnian and Lance L.P. Gore (London: Routledge, 2015), 193–215.

Christian Göbel - 9781786433787
Downloaded from Elgar Online at 10/28/2019 07:38:30PM
via free access

21. Andrew Kachites McCallum, "Mallet: A Machine Learning for Language Toolkit" (2002), http://mallet.cs.umass.edu.
22. For excellent introductions to topic modeling in political science, see Christopher Lucas et al., "Computer-Assisted Text Analysis for Comparative Politics," *Political Analysis* 23, no. 2 (2015): 254–77; Margaret E. Roberts, Brandon M. Stewart and Dustin Tingley, "Navigating the Local Modes of Big Data: The Case of Topic Models," in *Computational Social Science: Discovery and Prediction*, ed., R. Michael Alvarez (Cambridge: Cambridge University Press, 2016), 51–97; Justin Grimmer and Brandon M. Stewart, "Text As Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts," *Political Analysis* 21, no. 3 (2013): 267–97. For a more technical introduction by the developers of the technique, see David M. Blei, Andrew Y. Ng and Michael I. Jordan, "Latent Dirichlet Allocation," *Journal of Machine Learning Research* 3 (2003): 993–1022.
23. Keung Wong et al., "Rural Migrant Workers in Urban China: Living a Marginalised Life," *International Journal of Social Welfare* 16, no. 1 (2007): 32–40; Chih-Jou Jay Chen, "Die Zunahme von Arbeitskonflikten in China: Ein Vergleich von ArbeiterInnenprotesten in Verschiedenen Sektoren," in *Arbeitskämpfe in China: Berichte von der Werkbank der Welt*, eds., Georg Egger, Daniel Fuchs, Thomas Immervoll and Lydia Steinmassl (Vienna: Promedia, 2013), 78–105.
24. Xiaolin Guo, "Land Expropriation and Rural Conflicts in China," *The China Quarterly* 166 (2001): 422–39; You-tien Hsing, *The Great Urban Transformation: Politics of Land and Property in China* (Oxford: Oxford University Press 2012).
25. Sally Sargeson, "Violence as Development: Land Expropriation and China's Urbanization," *Journal of Peasant Studies* 40, no. 6 (2013): 1063–85; Kathy Le Mons Walker, "From Covert to Overt: Everyday Peasant Politics in China and the Implications for Transnational Agrarian Movements," *Journal of Agrarian Change* 8, no. 2–3 (2008): 462–88.
26. Kuang and Göbel, "Sustaining Collective Action in Urbanizing China."
27. On conflicts related to property management companies, see Luigi Tomba, *The Government Next Door: Neighborhood Politics in Urban China* (Ithaca, NY: Cornell University Press, 2014); Ying Wu and Junhua Chen, "The Constructive Significance of Homeowners' Rightful Protest in China," in *Neighbourhood Governance in Urban China*, ed., Ngai-Ming Yip (Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, 2014); Yongshun Cai, "China's Moderate Middle Class: The Case of Homeowners' Resistance," *Asian Survey* 45, no. 5 (2005): 777–99.
28. See for example H. Christoph Steinhardt and Fengshi Wu, "In the Name of the Public: Environmental Protest and the Changing Landscape of Popular Contention in China," *The China Journal* 75, no. 1 (2016): 61–82; Anna Lora-Wainwright and Benjamin Van Roij, "Learning to Live with Pollution: The Making of Environmental Subjects in a Chinese Industrialized Village," *The China Journal* 68 (2012): 106–24; Lei Xie, "China's Environmental Activism in the Age of Globalization," *Asian Politics & Policy* 3, no. 2 (2011): 207–24; Benjamin Van Rooij, "The People vs. Pollution: Understanding Citizen Action against Pollution in China," *Journal of Contemporary China* 19, no. 63 (2010): 55–77. Phillip Stalley and Dongning Yang, "An Emerging Environmental Movement in China?," *The China Quarterly* 186 (2006): 333–56; Jun Jing, "Environmental Protests in Rural China," in *Chinese Society: Change, Conflict and Resistance*, eds., Elizabeth J. Perry and Mark Selden (London and New York: Routledge, 2003), Chapter 9.
29. Graeme Lang and Ying Xu, "Anti-incinerator Campaigns and the Evolution of Protest Politics in China," *Environmental Politics* 22, no. 5 (2013): 832–48; see also Wedeman, "Not in My Backyard."
30. Steinhardt and Wu, "In the Name of the Public."
31. Thomas P. Bernstein and Xiaobo Lü, *Taxation without Representation in Contemporary Rural China* (Cambridge: Cambridge University Press, 2003).
32. Gerard A. Postiglione, *Education and Social Change in China: Inequality in a Market Economy* (Abingdon: Routledge, 2015).
33. On the diffusion of protests across state boundaries, see Stuart Hill and Donald Rothchild, "The Contagion of Political Conflict in Africa and the World," *Journal of Conflict Resolution* 30, no. 4 (1986): 716–35.
34. Elfstrom and Kuruvilla, "The Changing Nature of Labor Unrest in China": 465–66.
35. https://www.nytimes.com/2017/02/20/world/asia/xinjiang-china-police-rallies.html?_r=0; https://www.hrw.org/news/2015/01/29/china-political-repression-high-mark.
36. The number of protests is somewhat correlated with population size—other things being equal, cities with larger populations are also more likely to experience protests. If population size is controlled for, however, very small prefectures that experience only one or two protests are highlighted as extremely prone to protests.
37. Sidney G. Tarrow, *Power in Movement: Social Movements and Contentious Politics* (Cambridge: Cambridge University Press, 2011).
38. See, for example, Mark Granovetter, "Threshold Models of Collective Behavior," *American Journal*

Christian Göbel - 9781786433787
Downloaded from Elgar Online at 10/28/2019 07:38:30PM
via free access

of Sociology 83, no. 6 (1978): 1420–43; Mark Granovetter and Roland Soong, "Threshold Models of Diffusion and Collective Behavior," *Journal of Mathematical Sociology* 9, no. 3 (1983): 165–79.

39. John D. McCarthy, Clark McPhail and Jackie Smith, "Images of Protest: Dimensions of Selection Bias in Media Coverage of Washington Demonstrations, 1982 and 1991," *American Sociological Review* 61, no. 3 (1996): 478–99; Clark McPhail and John McCarthy, "Who Counts and How: Estimating the Size of Protests," *Contexts* 3, no. 3 (2004): 12–18; Mike Raybould et al., "Counting the Herd. Using Aerial Photography to Estimate Attendance at Open Events," *Event Management* 6, no. 1 (2000): 25–32.
40. On the relationship between different kinds of protests and state responses, and protests being "culturally sanctioned" in China, see Perry, *Challenging the Mandate of Heaven* and Elizabeth J. Perry, "Challenging the Mandate of Heaven: Popular Protest in Modern China," *Critical Asian Studies* 33, no. 2 (2001): 163–80.
41. Ching Kwan Lee, "Pathways of Labor Insurgency," in *Chinese Society: Change, Conflict and Resistance*, eds., Elizabeth J. Perry and Mark Selden (London and New York: Routledge, 2000), 73–95, 61; and Anita Chan, "Recent Trends in Chinese Labour Issues: Signs of Change," *China Perspectives*, no. 57 (2005): 1–13.
42. Zhou Guanghui, "Towards Good Government: Thirty Years of Administrative Reforms in China," in *The Reform of Governance*, ed., Keping Yu (Leiden: Brill, 2010), 137–80.
43. Barrington Moore, *Social Origins of Dictatorship and Democracy: Lord and Peasant in the Making of the Modern World* (Boston: Beacon Press, 1993); Dietrich Rueschemeyer, Evelyne Huber Stephens and John D. Stephens, *Capitalist Development and Democracy* (Chicago: University of Chicago Press, 1992).
44. I used a pipeline of Tfidf vectorizer, Tfidf transformer and Linear Support Vector Classification, a standard combination that was obtained from the machine learning library scikit-learn; see Fabian Pedregosa et al., "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research* 12 (2011): 2825–30.
45. The literature is far too extensive to be mentioned here. A comprehensive overview is provided in Kuang and Göbel, "Sustaining Collective Action in Urbanizing China" and Göbel and Ong, "Social Unrest in China."

Christian Göbel - 9781786433787
Downloaded from Elgar Online at 10/28/2019 07:38:30PM
via free access