Chinese Sociological Review

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/mcsa20

Does the crying baby always get the milk? An analysis of government responses for online requests

Tianji Cai, Fumin Li, Jian Zhan & Zhengrong Wang

To cite this article: Tianji Cai, Fumin Li, Jian Zhan & Zhengrong Wang (2022): Does the crying baby always get the milk? An analysis of government responses for online requests, Chinese Sociological Review, DOI: 10.1080/21620555.2022.2103667

To link to this article: https://doi.org/10.1080/21620555.2022.2103667

Published online: 05 Aug 2022.

CHINESE SOCIOLOGICAL REVIEW
https://doi.org/10.1080/21620555.2022.2103667

Does the crying baby always get the milk? An analysis of government responses for online requests

Tianji Cai[a], Fumin Li[a], Jian Zhan[b] and Zhengrong Wang[c]

[a]University of Macau, Macau, China; [b]Lanzhou University, Lanzhou, China; [c]Gansu Institute of Political Science and Law, Lanzhou, China

ABSTRACT

One thing that has frequently been overlooked in studies of how a local government responds to its citizens is the text itself. When a petitioner drafts a post, they can pick up words and choose how the demand or question is conveyed. Thus, how a post is composed may also influence local government responsiveness. In this study, we investigated whether a delayed response is due to the way the message is drafted or the actual content by separating content-related and content-agnostic text features. Based on posts retrieved from two websites, we found that the response pattern varies by location, time, and type of queried agency. Our results also indicated that lengthy and low positive sentiment posts generally result in longer waiting times. However, more research is needed to gauge the possibility of a false negative and the meaning of contents extracted from computational tools. In addition, our work scrutinized several methodological issues and sought practical solutions to analyzing big data.

Introduction

Regardless of regime type, both democratic and authoritarian governments respond to public demands, though their motivations may differ (e.g., Reilly 2012). However, a government's response can be very

CONTACT Zhengrong Wang wzr6673@gsupl.edu.cn School of Public Administration, Gansu Institute of Political Science and Law, Lanzhou, China.

selective. A petitioner's characteristics and policy domain, are linked to how the government responds to the request. As many government services are delivered online nowadays, and as web-based forums are increasingly used to collect public opinions, many social scientists are interested in how public opinions are responded to in an online context (Jiang, Meng, and Zhang 2019).

Over the past few decades, the Chinese government has invested resources to build its E-government network for data processing, information management, and transactions through government-funded projects. Online engagement between local government representatives and citizens has become a popular channel for individuals to raise concerns, ask questions, and express opinions on local affairs in China (Chen, Pan, and Xu 2016). The establishment of online forums to collect public opinion has attracted substantial scholarly attention. Taking advantage of web scraping, investigations that utilize text analysis techniques to study government responsiveness in China have been thriving (e.g., Su and Meng 2016). For example, many studies have explored the effects of institutional or socioeconomic factors at the macro level on government response; however, one thing that has frequently been overlooked is the text itself. When a petitioner drafts a request, he or she can choose how the demand or question is conveyed. Thus, how a post is composed rather than the content of the inquiry may also influence local government responsiveness. One may wonder if a poorly drafted message will result in a delayed response. It is still unclear whether a delayed response could be due to the way the message is drafted or what it is about. To fill this gap, this study explored the effect of text features on government responsiveness. In particular, we separated the content-related and the content-agnostic text features and investigated their effects on the speed of local government responses, accordingly. The content-related text features referred to what a post is about, such as policy domain (e.g., welfare and urban development), and type of inquiry (e.g., complaint and suggestion), while the content-agnostic text features covered those characteristics that are irrelevant to the message's content, such as its format (e.g., length and structure), sentiment (e.g., positive *vs.* negative), and text cohesion (e.g., connectives and semantic similarity).

The remainder of the article is organized as follows: First, we summarize previous studies on selective response and provide a general background of selective responses, how to quantify text features and the development of E-government in China. Then, a description of the data and methodology is presented, followed by the results and discussion. We conclude with a summary of the findings and suggestions for future studies.

Background

Selective responses

Previous studies have suggested that most local governments respond to citizens' requests, though their motivations for doing so may vary. For instance, the persistent pressure of winning elections is the major incentive for local officials to respond to public requests and deliver services in democratic societies (Broockman 2013); while in authoritarian regimes, such as China, the incentives for local officers to listen and respond to citizens' requests mainly originate from the considerations of maintaining durability (Gandhi 2008), such as oversight from the top and pressure from the bottom (Chen 2012), and providing public goods (Wang and Yao 2007).

Multiple studies have documented that the petitioner's characteristics (e.g., socioeconomic status, and race), the policy domain of the request (e.g., healthcare and labor), and institutional factors (e.g., level of government) are linked to three basic dimensions of responsiveness—speed, quality, and service orientation. Probably due to their constant participation and contributions to political candidates, affluent citizens are more likely to receive a response; and to see their preferences reflected in actual policy outcomes (Bartels 2006; Gilens 2005). Racial and ethnic discrimination in responsiveness has also been reported (Grohs, Adam, and Knill 2016). For instance, a field experiment revealed that emails with Black aliases' requests for voting registration receive fewer replies. Regardless of party affiliation, White legislators tend to be less responsive to Black aliases; while minority legislators show the opposite, with a higher level of responsiveness to Black alias (Butler and Broockman 2011). A similar field experiment focusing on religious discrimination conducted in China reported a mixed picture, with local governments in high-minority regions giving equal (if not privileged) responses to citizens with ethnic Muslim names. In contrast, local officials are less likely to help citizens with Muslim names in regions where the proportion of the Muslim population is low (Distelhorst and Hou 2014).

Besides petitioners' individual characteristics, policy domains are also important. Given limited public resources, a government must prioritize which issues need to be addressed first. In China, since the launch of the open-door policy, economic growth has been the focal point for the Chinese government (Jain 2017). Resources are prioritized toward policies that can promote economic growth, and public demands that can lead to potential economic growth are more likely to receive a response (Su and Meng 2016). Meanwhile, local government officials are less responsive to social welfare requests (Meng, Yang, and Su 2015). In addition, the petitioner's residence, the representation of demands, the degree of institutionalization, and the local leader's demonstration behavior, are relevant to the chance and speed of response in China (Distelhorst and Hou 2014).

Why text features matter

When an online request is submitted, it will be read and processed by a governmental officer, third-party evaluator, or even a contractor, all of whom have limited time to spend on each request. Thus, how a request is drafted could influence how it is perceived and handled because more urgent issues or better-formulated questions can be perceived and processed faster. The effect of wording has long been noticed in almost all social science disciplines and thus has attracted research attention from various fields, such as political science (e.g., Binder, Childers, and Johnson 2015), and survey studies (e.g., Weijters, Geuens, and Schillewaert 2010). Much attention has been paid to elaborating factors that influence comprehension in education and linguistic studies (Benjamin 2012). Referred to as readability, which indicates how difficult a text is perceived, studies have shown that both the complexity of the content (e.g., vocabulary and syntax) and typographic presentation of a text are related to its readability (Loyd 2013). Leaving out the effects of typography, readability theoretically also includes three linguistic features—lexical sophistication, syntactic complexity, and discourse structure (Snow 2002). In practice, measures of text-based features, such as the number of words per sentence, and the number of characters per word, have been widely adopted to assess readability (Crossley, Greenfield, and McNamara 2008), though the level of construct validity could be low, or a theoretical rationale could be missing (Norris and Ortega 2009).

Recent breakthroughs in natural language processing (NLP) have offered new measures closely related to the three linguistic features of readability (Kyle 2016). For instance, based on a word corpus assembled from a variety of sources, the study suggested that word frequency and word range can be used to measure lexical sophistication, and their results showed that both factors are important indicators of lexical and speaking proficiencies (Kyle and Crossley 2015). To quantify the syntactic complexity of a text, which can be broadly defined as the variation and sophistication of grammatical structures, measures at the text, paragraph, sentence, clause, and phrase levels have been proposed, such as the average length of sentences and clauses (Norris and Ortega 2009), the number of the noun and verb phrases (Biber, Gray, and Poonpon 2011), and the verb-argument constructions (Kyle 2016). Text cohesion, constructed from linguistic linkages in a text, has been commonly used as a proxy measure of discourse structures (Givón 1995). NLP tools nowadays can offer better estimates at the paragraph or sentence level for connectives, lexical overlap, and semantic similarity (Crossley, Kyle, and McNamara 2016), which allows researchers to differentiate global and local cohesions (Guo, Crossley, and McNamara 2013).

In addition, NLP tools can also be used to automatically summarize text contents and identify patterns (Nazari and Mahdavi 2019), which offers a practical way to quantify a text's content aside from its format. Thus, one could calibrate the effect of how a text is drafted apart from what it is about (Tan, Lee, and Pang 2014). For instance, studies have reported that spelling errors (Askira-Gelman and Barletta 2008), emotions (Berger and Milkman 2012), and informativeness (Dilip et al. 2018) are related to the quality or popularity of text controlling the content-related measures, such as topics.

The majority of the existing studies and tools are available only in English, with a handful of research conducted in other languages, particularly Chinese (Sung, Chang, and Yang 2015). With new open source tools and platforms (e.g., PaddlePaddle by Ma et al. 2019), many indices that assess lexical, syntactic, and discourse features, as well as NLP models that extract content-related and content-agnostic measures, have become available for use with Chinese text.

Development of E-government in China

E-government's emergence is an opportunity to reshape the public sector by reducing bureaucratic costs and improving relationships between citizens and the government in China (Zhang 2001). Following the advice of the *Government Online* program (Government Online Project 2000), most governmental agencies, such as those in cities at the prefecture level and above, as well as ministries and commissions, began to build an online platform upon which to disclose information and provide public service announcements. In 2008, President Hu Jintao called for more proactive government measures for responding to and consulting public opinion and advocated for establishing channels that would bridge government and citizens (Jia 2019). Thus, besides releasing information and providing public services, government bodies have invested in sites and platforms for canvassing public opinion and incorporating it into the policymaking process, which has been acknowledged as a low-cost-high-efficiency measure to improve capacity for governance (State Council 2016). Microblogs, emails, message boards, and online surveys became utilized to collect complaints and appeals, solve problems, and interact with citizens (Esarey 2015; Schlaeger and Jiang 2014).

Establishing online forums to collect public opinion in China has cultivated research interests that utilize automated text analysis techniques to study how the government responds to online requests. Studies have identified that content-related factors, such as policy issues and the type of request, as well as certain text features, are relevant to the chance and the speed of a response (e.g., Su and Meng 2016). However, a systematic evaluation of text features that include content-related and content-agnostic measures is still lacking. For example, one may wonder whether a poorly drafted post (e.g., badly organized, full of spelling errors) or an ambiguous post without a clear indication of the subject will delay a response. In contrast, posts with a positive tone or strong emotion may receive a prompt response.

Therefore, this study had two objectives: first, to evaluate the effects of both content-related and content-agnostic text features on government responsiveness; second, to explore practical solutions for integrating the results obtained from data mining tools in social science studies. In particular, we hypothesized that a better-written inquiry (e.g., concise, fluent, and positive) would be more likely to get a faster response holding everything else, such as the petitioner's characteristics, policy domain, and institutional factors, being equal.

Data and methods

Data

The current study retrieved all posts on the websites of Luzhou Wenzheng (泸州问政)[1] and Chengdu Lizheng Mayor's Mailbox (成都网络理政市长信箱),[2] which are both official online platforms for collecting public opinions. Luzhou City is a prefecture-level city in Sichuan province with a population of 4.5 million, of which ~25% of the residents live in the metro area. Chengdu is the capital of Sichuan province with 16.6 million residents living in twelve districts, five prefecture-level cities, and three counties (Chengdu Yearbook Society 2019).

Although most local governments set up message boards after 2010 to respond to the central government's requirements, many of them lost their popularity as other types of social media tools became available. Probably due to the local leader's personal effort of improving governmental responsiveness, Luzhou Wenzheng is one of the exceptions. The website has become a popular channel for locals to express opinions and send requests since its birth in May 2012, with an average of more than 1,200 posts per month in 2020. The website is not limited to local users as long as a valid phone number is provided for registration. However, since users must choose an appropriate local agency while submitting, requests must be relevant to local businesses. All available posts—a total of 82,474 (05/19/2012–10/11/2020) were retrieved from the website.

Unlike the Luzhou Wenzheng website, which is the city's only official channel, due to the diversity and volume of requests received each day, the website of the Chengdu Municipal government offers multiple channels to collect public opinions. For example, Mayor's Mailbox aims to collect constructive opinions on issues related to administration, service, development, environment, and welfare; while questions and complaints that fall into a specific division of governmental responsibilities are handled directly by their own mailbox systems. To simplify our study, we only focused on Mayor's Mailbox and fetched all 14,756 posts accessible (01/17/2014–09/01/2020). The primary reason for choosing Mayor's Mailbox in Chengdu instead of other cities was because of its data structure. Both Luzhou and Chengdu share similar data structures in terms of types of agencies, time of launch, and distribution of response time. Therefore, Chengdu was used to train our models, and the results could serve as a replication and robustness check of our model specifications.

Variables

We chose the speed of government response, measured by the number of working days between the timestamp of the request and that of the response, as our dependent variable to investigate responsiveness. The dependent variable ranged from 0 to 190 for the Luzhou data and from 0 to 197 for the Chengdu data, respectively.[3] Two distributional features of the dependent variable needed to be addressed in the modeling process. First, the number of working days for both datasets showed distinct tail distribution patterns (about 95% of the cases had a value <14), and the remaining 5% appeared to be highly skewed to the right. Secondly, multiple inflated points require suitable models beyond a traditional count regression (Cai, Xia, and Zhou 2021). As listed in their performance pledges, the Chengdu government will respond to regular requests within five working days; for complicated requests, the number of needed working days could be extended to 10 and 20, respectively. Thus, one would expect heaps at values of 5, 10, and 20 on the dependent variable for Chengdu.

Measures of text features

We utilized NLP tools from the PaddlePaddle platform to calculate the content-related and content-agnostic measures for the retrieved posts. For example, we applied the Latent Dirichlet Allocation (LDA) model (Blei, Andrew, and Jordan 2003) to extract topics from each post as one of the content-related measures. LDA model estimates the probability that a new text belongs to a given set of topics characterized by keywords obtained from the corpus defined by a set of documents. Unlike the traditional LDA model, which requires specifying the number of topics in advance, PaddlePaddle offers Familia LDA models (Jiang et al. 2018) that are pre-trained from three large-scale industrial sets of corpora with an optimized number of topics: Baidu news in 2016 (2,000 topics), Baidu novel dataset (500 topics), and Baidu webpage dataset (4,267 topics), accordingly. Thus, utilizing the Familia LDA models could reduce the bias or error resulting from human interaction (e.g., manually removing stop words and general words, and tuning the model parameters), because the pre-trained topics were obtained from comprehensive corpora that are less dependent on specific text corpus or individual researcher. Besides computational efficiency, utilizing the version-traceable pre-trained model also made our results reproducible, since essentially all users are calling the same pre-trained model.

However, the resulting probabilities were vast in size and challenging to interpret (e.g., 2,000 topics for each post based on the news corpus). Due to its size, it was difficult to include them in any statistical modeling without further dimensionality reductions. Moreover, because the probabilities for each post summed to one, the resulting matrix was subject to unit-sum constraints. Thus, we first applied center log-ratio transformation, in which zero values were treated as being below the detection limit, to avoid the potential negative bias (Van den Boogaart and Tolosana-Delgado 2013). Principal component analysis (PCA) was then conducted, followed by variable selection among the resulting PCs using the dependent variable as the target under a negative binomial assumption. The selected PCs explained about 25.2% and 33.6% of the variation for the covariance of the transformed LDA results for the Luzhou data and Chengdu data, respectively. Finally, the selected PCs were similarly enclosed in statistical models to control population structure in behavioral genetics (Crosslin et al. 2014).[4]

In terms of the content-agnostic measures, the Jensen-Shannon divergence (JSD) (Lin 1991) was used to estimate the correlation between two adjacent sentences within each of the posts. JSD is one of the most frequently-used measures to quantify divergence between two corpora (Koplenig 2017). Ranging from 0 to 1, it measures the distributional similarity (e.g., keywords and their probabilities) of two topic corpora, and a higher value suggests the corpora have more distinguishing words. Since the number of sentences varies dramatically from one post to another, using centrality measures, such as the mathematic average may introduce systematic bias, and the related variability measures, such as the variance or standard deviation, could also provide an inflated or biased estimate of variability (Wiley et al. 2014). Therefore, we chose to use the median (Med JSD) and Median

Absolute Difference of JSD (MAD JSD) to quantify the average level of fluency and its variability. Since JSD does not tell whether any two sentences were grammatically correct, it was possible that none of them would be grammatically fluent but have similar topic structures. Unfortunately, cohesion indices that are more suitable for quantifying the level of fluency, such as lexical overlap and syntactic cohesion (Crossley, Kyle, and McNamara 2016), are not available for the Chinese language. Therefore, we included the number of word-level typos and character errors in a post as a proxy for the global level of cohesion. The number of errors in a post (Err) was estimated using the *pycorrector* package to detect and correct errors in Chinese (Xu 2022).

To measure the tone of each post, we adopted the results obtained from the sentiment analysis of PaddlePaddle, which was based on the sentiment knowledge enhanced pretraining model, to provide a unified sentiment representation for multiple sentiment analysis tasks (Tian et al. 2020). The results contained a measure of the overall positivity, a measure of the positivity of each sentence, and the classification (positive *vs.* negative) of each sentence. Similar to the JSD measures, we calculated the median (Med Sent) and MAD of the positivity (MAD Sent), as well as the proportion of positive classifications (PropS). In addition, logarithms of the number of words (LogW), the number of topics (LogT), and average words per sentence (LAWS) were also estimated.

The indices of the text features could be divided into the content-related and the content-agnostic groups. For instance, content-related indices, such as the PCs of the topics were used to control the contents' effect. In contrast, the content-agnostic indices evaluated the syntactic complexity (e.g., LogW, LAWS, LogT, and Err), the sentiment (e.g., Positive Sent., PropS, Med Sent, and MAD Sent), and the cohesion (e.g., Med JSD and MAD JSD) of a post. In line with previous studies (e.g., Su and Meng 2016), we also controlled for the year of the posts, whether the petitioner used a real name or phone number (for the Luzhou data only), type of queried agency, and level of agency.

As shown in Figure 1, for each dataset, duplicated posts, posts that did not receive a response, or those that contained missing values on the measures of text features were removed. For example, among the 82,451 posts retrieved from Luzhou Wenzheng, 314 cases had missing values for the dependent variable (no response at the moment of retrieving), and 3,382 cases did not have measures of the similarity between sentences (e.g., JSD) because those posts only contained one sentence. The final sample size for the Luzhou and Chengdu data was 78,738 and 14,109, respectively.

Methods

When data are characterized by a highly skewed long tail and overdispersion, neither the Poisson nor the negative binomial model is adequate (Gupta and Ong 2005). Many models have been proposed under the framework of mixed Poisson distributions to address the long-tail issue. In general, a mixed Poisson distribution can be formulated as follows:

$$P(k) = \int_0^\infty e^{-\lambda} \frac{\lambda^k}{k!} g(\lambda) d\lambda$$

where the Poisson mean (λ) is a random variable with the probability density function $g(\lambda)$. We picked the inverse Gaussian distribution for $g(\lambda)$ that yields the Poisson inverse Gaussian (PiG) model (Willmot 1987). In addition, a generalized negative binomial (GNB) model (Gupta and Ong 2004) was also implemented. For the Chengdu data, inflated versions of the negative binomial (NB) and GNB models were also explored (Cai, Xia, and Zhou 2021). The PiG model was estimated using the R library's *gamlss* (Stasinopoulos and Rigby 2007), and the rest were implemented in the SAS NLMIXED procedure (High 2017).

While large sample size is always desirable because it reduces the occurrence of type II error, it also amplifies the detection of trivial differences that are not substantially meaningful. As one of the issues related to big data analysis, running statistical models on data with a size $>$10,000 inevitably increases the chance of a false positive occurring (Kaplan, Chambers, and Glasgow 2014). To alleviate such a problem, except for descriptive statistics, we reported coefficients estimated from the whole sample but using 95% of bootstrapping confidence intervals (BCIs) obtained from 500 bootstrapping replicates with a size of 5,000 each.

Results

Descriptive analysis

Figure 2 shows the number of posts, average working days to receive a response and the number of posts per month over the years for both data sets. The bar represents the average working days to receive a response by month, the error bar indicates one-tenth of the standard deviation, and the solid line refers to the number of posts per month. There was a noticeable regional gap between Luzhou and Chengdu for the trend of the posts and responses. According to panel A in Figure 2, the number of posts per month for Luzhou kept an increasing trend after the platform was launched in May 2012 and showed visible fluctuations, usually during the month of Chinese New Year, and the average number of posts per

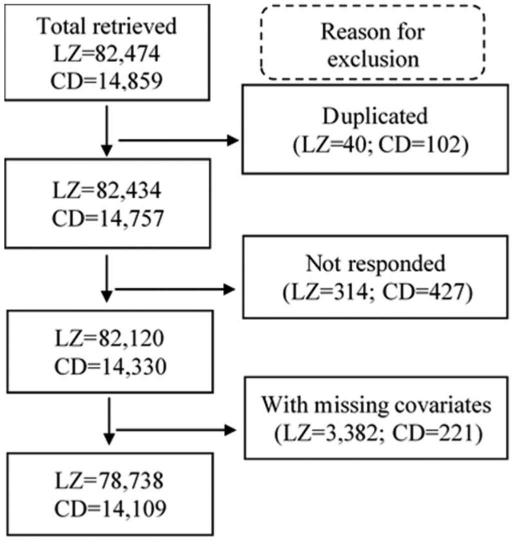

Figure 1. Final sample size flowchart.

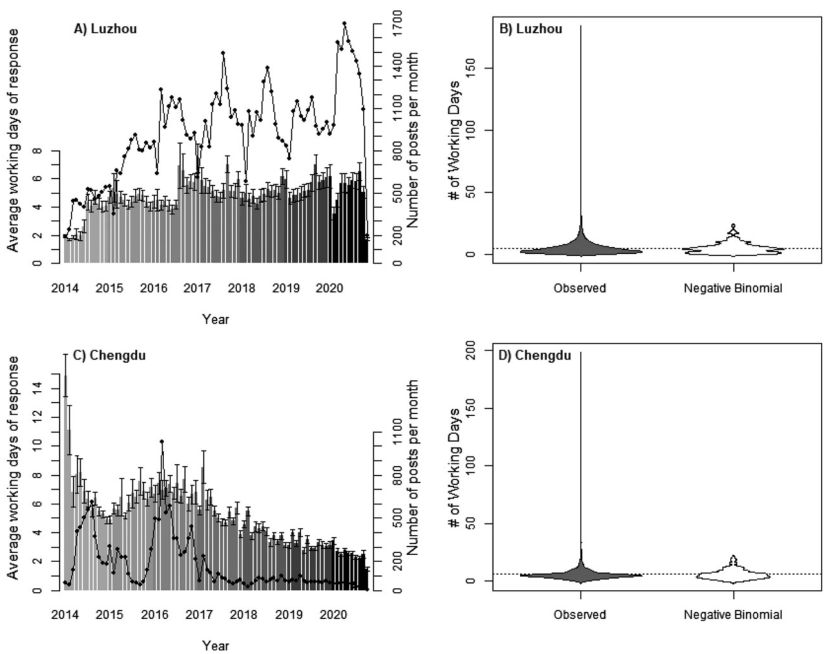

Figure 2. Distribution of the number of working days to receive a response for the Luzhou and Chengdu datasets.

month reached more than 1,000 after 2015; meanwhile, the average of working days to receive a response increased after 2014, from lower than four days to more than five days. As shown in panel C, the trend of Chengdu experienced some dramatic ups and downs from 2014 to the first half of 2017, with a peak in March 2016, and then followed a stable declining path, with the number of monthly posts reaching as low as 50. The average number of working days dropped from more than six days to lower than four days after 2017.

Panels B and D compare the empirical distribution of the number of working days and the theoretical distribution based on a negative binomial assumption according to the parameters estimated from the observed data. For both places shown in the bean chart, the dependent variable exhibited a heavy-tailed distribution that was beyond the capability of the regular negative binomial model for handling over-dispersion. Thus, it was necessary to employ models that were suitable for modeling such features.

Tables 1 and 2 report the descriptive statistics for the variables included in the analysis of Luzhou and Chengdu, respectively. Consistent with the trend shown in Figure 2, the average working day to receive a response increased from 1.92 to 5.42 days for Luzhou. In 2014, the local government in Luzhou revised the website to allow petitioners to specify which governmental agency to send the request to. By the policy domain that posts related to (Su and Meng 2016), we classified the posts according to their queried 419 agencies into eight groups: "Public transportation," "Public safety," "Urban construction," "Welfare," "Planning," "Administration and law enforcement," "Agricultural, cultural, communicational and recreational issues," and "Other" that covered those that did not or hard to fit a specific governmental agency. Except for the "other" group, welfare, agricultural issues, and urban construction were the top three requested groups, and the percentage of "other" requests declined over time, especially after 2017.

Regarding the administrative level of the queried agencies, about two-thirds of posts requested prefectural level ones. The popularity of agricultural-related requests and the high proportion of prefectural level agencies may relate to the large size of Luzhou's population living in non-metro areas. In terms of the text characteristics, a typical post was about 100 words [$\exp(4.61) = 100.5$] long and contained 16 words per sentence [$\exp(2.76) = 15.8$], with a total of 16 topics [$\exp(2.79) = 16.3$] identified from the LDA model based on the news topic set. All three characteristics were relatively stable, with slight variations over time. Users published about a quarter of the posts possibly using their real name IDs or phone numbers ($0.17 + 0.06 = 0.23$). Since this study relied on a publicly-available database to identify real names and phone numbers, no

Table 1. Descriptive statistics for Luzhou data by year of post.

	Category	Total Mean (*SD*)	2012 Mean (*SD*)	2013 Mean (*SD*)	2014 Mean (*SD*)	2015 Mean (*SD*)	2016 Mean (*SD*)	2017 Mean (*SD*)	2018 Mean (*SD*)	2019 Mean (*SD*)	2020 Mean (*SD*)
# of WD		4.99 (5.67)	1.92 (0.88)	1.97 (4.40)	3.32 (3.96)	4.59 (5.01)	5.00 (7.40)	5.53 (5.65)	4.95 (4.67)	5.49 (5.34)	5.42 (5.93)
Level of agency	City	0.35 (0.48)	–	–	0.20 (0.40)	0.41 (0.49)	0.41 (0.49)	0.38 (0.49)	0.36 (0.48)	0.33 (0.47)	0.33 (0.47)
	Prefectural	0.65 (0.48)	–	–	0.80 (0.40)	0.59 (0.49)	0.59 (0.49)	0.62 (0.49)	0.64 (0.48)	0.67 (0.47)	0.67 (0.47)
Queried agency	Public transp.	0.06 (0.23)	–	–	0.03 (0.18)	0.06 (0.23)	0.05 (0.22)	0.04 (0.20)	0.06 (0.24)	0.07 (0.26)	0.07 (0.26)
	Public safety	0.09 (0.29)	–	–	0.03 (0.18)	0.08 (0.26)	0.10 (0.30)	0.10 (0.30)	0.10 (0.30)	0.11 (0.31)	0.10 (0.31)
	Urban const.	0.10 (0.29)	–	–	0.05 (0.21)	0.05 (0.21)	0.05 (0.23)	0.10 (0.30)	0.17 (0.37)	0.12 (0.33)	0.11 (0.31)
	Welfare	0.11 (0.31)	–	–	0.04 (0.19)	0.08 (0.26)	0.08 (0.26)	0.10 (0.30)	0.13 (0.34)	0.16 (0.37)	0.16 (0.36)
	Planning	0.05 (0.21)	–	–	0.01 (0.11)	0.05 (0.22)	0.05 (0.21)	0.05 (0.23)	0.05 (0.23)	0.05 (0.22)	0.04 (0.20)
	Adm./Law	0.07 (0.26)	–	–	0.01 (0.10)	0.04 (0.20)	0.04 (0.19)	0.07 (0.25)	0.09 (0.28)	0.12 (0.33)	0.11 (0.31)
	Agri./Cul./Rec.	0.11 (0.31)	–	–	0.03 (0.18)	0.07 (0.26)	0.06 (0.23)	0.08 (0.27)	0.13 (0.33)	0.15 (0.36)	0.18 (0.38)
	Other	0.42 (0.49)	–	–	0.79 (0.40)	0.58 (0.49)	0.58 (0.49)	0.46 (0.50)	0.27 (0.45)	0.21 (0.40)	0.23 (0.42)
LogT		2.79 (0.37)	2.85 (0.31)	2.83 (0.35)	2.83 (0.36)	2.81 (0.37)	2.80 (0.37)	2.80 (0.37)	2.77 (0.37)	2.78 (0.37)	2.78 (0.37)
LogW		4.61 (0.79)	4.60 (0.79)	4.67 (0.80)	4.72 (0.85)	4.69 (0.82)	4.67 (0.80)	4.64 (0.79)	4.55 (0.77)	4.57 (0.77)	4.53 (0.77)
Positive Sent		0.30 (0.29)	0.37 (0.33)	0.31 (0.31)	0.31 (0.30)	0.31 (0.29)	0.32 (0.29)	0.32 (0.29)	0.30 (0.28)	0.28 (0.28)	0.28 (0.28)
PropS		0.43 (0.24)	0.46 (0.26)	0.43 (0.24)	0.43 (0.24)	0.43 (0.24)	0.43 (0.24)	0.44 (0.24)	0.43 (0.25)	0.41 (0.25)	0.41 (0.25)
Err		0.03 (0.16)	0.03 (0.16)	0.03 (0.17)	0.04 (0.19)	0.03 (0.18)	0.03 (0.17)	0.03 (0.17)	0.02 (0.15)	0.02 (0.15)	0.02 (0.15)
Post ID	Real name	0.17 (0.37)	0.36 (0.48)	0.30 (0.46)	0.25 (0.43)	0.11 (0.31)	0.13 (0.34)	0.13 (0.33)	0.17 (0.38)	0.19 (0.39)	0.19 (0.40)
	Phone #	0.06 (0.25)	0.00 (0.00)	0.00 (0.00)	0.02 (0.14)	0.05 (0.21)	0.06 (0.24)	0.08 (0.27)	0.06 (0.23)	0.07 (0.26)	0.09 (0.28)
	Other	0.77 (0.42)	0.64 (0.48)	0.70 (0.46)	0.73 (0.44)	0.85 (0.36)	0.81 (0.39)	0.79 (0.41)	0.77 (0.42)	0.74 (0.44)	0.72 (0.45)
LAWS		2.76 (0.34)	2.80 (0.34)	2.76 (0.32)	2.76 (0.34)	2.75 (0.33)	2.75 (0.33)	2.76 (0.34)	2.77 (0.36)	2.77 (0.35)	2.77 (0.35)
Med Sent.		0.39 (0.23)	0.43 (0.24)	0.40 (0.23)	0.40 (0.23)	0.40 (0.22)	0.40 (0.22)	0.41 (0.23)	0.40 (0.23)	0.38 (0.23)	0.38 (0.23)
MAD Sent.		0.16 (0.09)	0.16 (0.10)	0.16 (0.09)	0.16 (0.09)	0.16 (0.09)	0.16 (0.09)	0.16 (0.09)	0.16 (0.09)	0.15 (0.10)	0.15 (0.10)
Med JSD		0.03 (0.09)	0.04 (0.09)	0.04 (0.09)	0.03 (0.09)	0.04 (0.09)	0.03 (0.09)	0.03 (0.09)	0.04 (0.09)	0.03 (0.09)	0.03 (0.09)
MAD JSD		0.01 (0.03)	0.01 (0.04)	0.01 (0.03)	0.01 (0.03)	0.01 (0.03)	0.01 (0.03)	0.01 (0.03)	0.01 (0.04)	0.01 (0.03)	0.01 (0.03)
N		78,738	258	1,901	5,159	8,824	11,980	12,581	12,246	11,963	13,826

Table 2. Descriptive statistics for Chengdu data by year of post.

	Category	Total Mean (*SD*)	2014 Mean (*SD*)	2015 Mean (*SD*)	2016 Mean (*SD*)	2017 Mean (*SD*)	2018 Mean (*SD*)	2019 Mean (*SD*)	2020 Mean (*SD*)
# of WD		6.21 (6.75)	6.69 (7.85)	5.99 (6.48)	6.98 (6.86)	6.06 (6.50)	3.95 (2.67)	3.31 (2.25)	2.66 (1.69)
Level of agency	City	0.86 (0.35)	0.83 (0.37)	0.82 (0.39)	0.86 (0.34)	0.93 (0.26)	0.91 (0.29)	0.89 (0.32)	0.86 (0.35)
	Prefectural city	0.05 (0.21)	0.05 (0.22)	0.04 (0.19)	0.04 (0.20)	0.05 (0.22)	0.06 (0.24)	0.08 (0.28)	0.08 (0.28)
	Prefecture	0.09 (0.29)	0.12 (0.32)	0.14 (0.35)	0.10 (0.30)	0.02 (0.15)	0.03 (0.17)	0.03 (0.17)	0.05 (0.23)
Queried agency	Public transp.	0.23 (0.42)	0.27 (0.44)	0.26 (0.44)	0.26 (0.44)	0.25 (0.43)	0.01 (0.11)	0.01 (0.08)	0.01 (0.11)
	Public safety	0.05 (0.22)	0.04 (0.21)	0.05 (0.22)	0.05 (0.22)	0.04 (0.20)	0.04 (0.20)	0.07 (0.25)	0.10 (0.30)
	Urban const.	0.15 (0.36)	0.14 (0.35)	0.16 (0.37)	0.13 (0.33)	0.11 (0.32)	0.30 (0.46)	0.24 (0.43)	0.21 (0.41)
	Welfare	0.07 (0.25)	0.05 (0.22)	0.07 (0.25)	0.07 (0.26)	0.06 (0.24)	0.12 (0.33)	0.10 (0.30)	0.08 (0.27)
	Planning	0.21 (0.41)	0.20 (0.40)	0.21 (0.40)	0.20 (0.40)	0.21 (0.41)	0.29 (0.45)	0.31 (0.46)	0.21 (0.41)
	Adm./Law	0.20 (0.40)	0.23 (0.42)	0.17 (0.38)	0.22 (0.41)	0.24 (0.43)	0.05 (0.21)	0.07 (0.25)	0.12 (0.33)
	Agri./Cul./Rec.	0.09 (0.29)	0.06 (0.24)	0.08 (0.27)	0.07 (0.25)	0.08 (0.26)	0.20 (0.40)	0.22 (0.41)	0.27 (0.44)
LogT		2.97 (0.31)	3.03 (0.28)	3.01 (0.29)	2.96 (0.31)	3.03 (0.28)	2.89 (0.35)	2.80 (0.38)	2.67 (0.35)
LogW		4.43 (0.81)	4.63 (0.78)	4.58 (0.77)	4.38 (0.79)	4.53 (0.78)	4.17 (0.76)	3.90 (0.72)	3.63 (0.58)
Positive Sent		0.26 (0.31)	0.25 (0.30)	0.26 (0.31)	0.23 (0.30)	0.28 (0.32)	0.35 (0.34)	0.30 (0.30)	0.32 (0.29)
PropS		0.03 (0.04)	0.04 (0.04)	0.04 (0.04)	0.03 (0.03)	0.04 (0.04)	0.03 (0.04)	0.02 (0.02)	0.01 (0.01)
Err		0.06 (0.23)	0.08 (0.26)	0.07 (0.25)	0.05 (0.23)	0.05 (0.23)	0.02 (0.14)	0.01 (0.08)	0.00 (0.05)
LAWS		2.03 (0.32)	2.04 (0.32)	2.04 (0.30)	2.04 (0.33)	2.03 (0.28)	2.01 (0.29)	1.98 (0.31)	2.04 (0.35)
Med Sent.		0.38 (0.24)	0.38 (0.23)	0.39 (0.24)	0.36 (0.24)	0.39 (0.25)	0.43 (0.25)	0.40 (0.23)	0.42 (0.23)
MAD Sent.		0.18 (0.10)	0.19 (0.10)	0.18 (0.10)	0.17 (0.10)	0.18 (0.10)	0.17 (0.10)	0.16 (0.10)	0.16 (0.09)
Med JSD		0.02 (0.07)	0.02 (0.06)	0.02 (0.06)	0.02 (0.07)	0.02 (0.06)	0.02 (0.07)	0.04 (0.09)	0.04 (0.09)
MAD JSD		0.01 (0.03)	0.01 (0.02)	0.01 (0.02)	0.01 (0.03)	0.01 (0.02)	0.01 (0.03)	0.01 (0.04)	0.01 (0.04)
N		14,109	3,728	1,938	5,389	1,137	804	757	356

validation was conducted, and the results may have contained substantial measurement errors. The average positivity for the posts was about .30, with a slightly declining trend over time from .37 in 2012 to .28 in 2020. The average proportion of sentences that were classified as being positive (positivity $\geq$ .5) was about .43, with slight variations over time, which was consistent with the measures of the median and MAD, both of which showed stable values and low variability of positivity for each sentence over time. The value of the median JSD (.03) indicated the average similarity between two consecutive sentences was high, and the MAD of JSD also presented a steady trend with low variations over time. Because the number of posts was relatively small and some independent variables were not available for 2012 and 2013, only posts after 2013 were used for the later analysis.

For the Chengdu-related data, the average working day to receive a response followed a declining trend. For example, probably due to the decline in the volume of posts over the years, it dropped from 6.69 days in 2014 to 2.66 days in 2020. The Mayor's Mailbox website requests petitioners to select the listed category to which the opinion belongs. Similarly, we regrouped the listed categories to be consistent with the ones used in Luzhou (except for the "other" category, which was not available for the Mayor's Mailbox at the time of retrieval). Overall, the top three categories of queried agencies were public transportation (23%), planning (21%), and administration and law enforcement (20%). However, from 2014 to 2020, opinions related to public transportation and administration and law enforcement dropped considerably, from 27 to 1% and 23 to 12%, respectively; while posts related to urban construction and agricultural, cultural, communicational, and recreational issues surged quickly, from 14 to 21% and 6 to 27%, accordingly.

Due to Chengdu's large urban area, 86% of opinions were at the city level, and those submitted to agencies at the prefectural city level and prefecture-level were small (5 and 9%, respectively). Compared to the average post in Luzhou, the average post in Chengdu was slightly shorter in words (83.9 *vs.* 100.5) and sentences (7.6 *vs.* 15.8 words) but covered more topics (19.5 *vs.* 16.3). The average positivity of the posts in Chengdu was lower than that in Luzhou (.26 *vs.* .30); however, the overall positivity increased from .25 in 2014 to .32 in 2020. Although the average proportion of sentences classified as being positive in each of the posts did not change much, the median positivity increased gradually along with a slight decline in variation, as indicated by the value of MAD over time. The increase in positivity may have been due to the selection effect. For instance, those whose posts were more negative stopped posting in the forum. The consistent value of the median and MAD of JSD also demonstrated that the sentences in each post were well-connected. Since the Chengdu data collection did not cover all of 2020 and the number of posts was much smaller than that in 2019, we combined the posts from 2019 and 2020 in the later analysis.

Effect of text features on the response

To evaluate how the local governments responded to online requests, a series of regression models using the number of working days to receive a reply as the dependent variable were estimated, as shown in Table 3. The GNB model and PiG model were implemented to address the heavy tail issue. In addition, to alleviate any false positives due to the large sample size, all coefficients were evaluated by empirical confidence intervals obtained from 500 bootstrapping replicates (sampled with replacement) with the size of 5,000.[5] Models 1 and 2 included a group of control variables, such as year of the post, type of queried agency, level of queried agency, and type of ID (real, phone number, and other), as well as content-agnostic measures of the syntactic complexity, sentiment, and cohesion. Built on models 1 and 2, models 3 and 4 added the selected PCs obtained from the LDA results to evaluate the effect of content-related measures on the dependent variable.

Consistent with the descriptive statistics, the year of the post had a significant impact on the days needed for response. Requests posted in 2014 and 2015 were responded to much faster than those posted in later years. For instance, according to model 1, the ratio of the expected number of working days needed for response between 2014 and 2020 was about .70 [$\exp(-0.361) = .70$], holding other covariates constant. The type of queried agency was also associated with the time needed across the four models. For example, posts passed to agencies related to welfare and agricultural, cultural, communicational, and recreational issues received a response more quickly. In contrast, posts sent to administration and law enforcement, and public safety agencies need more days for a response. One possible reason for the discrepancy is that the former is more about existing policies while the latter might need time to investigate or collect information. In terms of the content-agonistic features, the length of a post was linked to a longer response time for all four models, e.g., positive coefficient 0.059 with 95% of BCI [0.017, 0.120] and 0.085 with 95% of BCI [0.039,0.125] in models 3 and 4, respectively. Although the sign of the coefficients and the model-based significance were consistent across the four models, the other content-agonistic features, such as the proportion of positive classifications and the number of topics, were not robust according to the bootstrapping procedure. For example, models 1 and 2 indicated that posts with a higher level of positive sentiment were associated with a shorter waiting time, negative coefficient −0.223 [−0.347,

Table 3. Estimated coefficients for the number of working days to receive a response for the Luzhou data.

Parameter	Category	Model 1 GNB Beta [2.5th, 97.5th]	Model 2 PiG Beta [2.5th, 97.5th]	Model 3 GNB w./PCs Beta [2.5th, 97.5th]	Model 4 PiG w./PCs Beta [2.5th, 97.5th]
Intercept		1.030 [0.650, 1.510]***#	0.946 [0.603, 1.249]***#	1.151 [0.813, 1.686]***#	1.148 [0.787, 1.470]***#
LogT		0.052 [0.031, 0.136]***#	0.103 [0.059, 0.142]***#	0.059 [0.017, 0.120]***#	0.085 [0.039, 0.125]***#
LogW		0.068 [−0.041, 0.175]***	0.071 [−0.015, 0.163]***	0.024 [−0.100, 0.128]**	0.020 [−0.071, 0.113]$^{+}$
Positive Sent		0.015 [−0.251, 0.268]	0.037 [−0.160, 0.252]	−0.014 [−0.306, 0.242]	0.008 [−0.205, 0.221]
PropS		−0.223 [−0.347, −0.053]***#	−0.225 [−0.347, −0.103]***#	−0.114 [−0.247, 0.070]***	−0.109 [−0.237, 0.016]***
Err		0.020 [−0.155, 0.242]	0.018 [−0.121, 0.188]	0.019 [−0.153, 0.239]	0.017 [−0.128, 0.188]
Post ID	Real name	−0.029 [−0.107, 0.056]***	−0.023 [−0.095, 0.042]**	−0.014 [−0.094, 0.076]$^{+}$	−0.007 [−0.078, 0.059]
	Phone #	0.022 [−0.110, 0.127]$^{+}$	0.003 [−0.094, 0.109]	0.037 [−0.100, 0.137]**	0.017 [−0.078, 0.117]
	Other				
LAWS		0.038 [−0.051, 0.132]***	0.044 [−0.024, 0.127]***	0.020 [−0.068, 0.117]*	0.027 [−0.046, 0.108]**
Med Sent.		−0.126 [−0.441, 0.180]***	−0.155 [−0.416, 0.084]***	−0.046 [−0.354, 0.256]	−0.066 [−0.332, 0.161]*
MAD Sent.		−0.044 [−0.365, 0.313]	−0.045 [−0.328, 0.264]	−0.016 [−0.331, 0.336]	−0.009 [−0.302, 0.284]
Med JSD		0.008 [−0.341, 0.308]	0.016 [−0.308, 0.318]	−0.066 [−0.409, 0.247]$^{+}$	−0.035 [−0.348, 0.262]
MAD JSD		0.209 [−0.682, 1.343]*	0.182 [−0.576, 1.040]$^{+}$	0.186 [−0.646, 1.315]*	0.163 [−0.575, 1.038]$^{+}$
Level of agency	City	0.094 [0.038, 0.234]***#	0.108 [0.032, 0.182]***#	0.063 [0.004, 0.193]***#	0.072 [−0.004, 0.145]***
	Prefectural	–	–	–	–
Queried agency	Public transp.	0.205 [−0.086, 0.246]***	0.191 [0.083, 0.316]***#	0.155 [−0.140, 0.209]***	0.133 [0.017, 0.256]***#
	Public safety	0.198 [−0.068, 0.244]***	0.163 [0.059, 0.274]***#	0.185 [−0.073, 0.247]***	0.148 [0.037, 0.267]***#
	Urban const.	−0.025 [−0.214, 0.051]*	−0.097 [−0.208, 0.014]***	−0.041 [−0.247, 0.039]**	−0.105 [−0.228, 0.018]***
	Welfare	−0.210 [−0.431, −0.175]***#	−0.252 [−0.356, −0.153]***#	−0.138 [−0.378, −0.091]***#	−0.186 [−0.297, −0.067]***#
	Planning	0.127 [−0.159, 0.181]***	0.113 [−0.005, 0.232]***	0.101 [−0.174, 0.157]***	0.090 [−0.029, 0.206]***
	Adm./Law	0.366 [0.189, 0.415]***#	0.273 [0.165, 0.370]***#	0.302 [0.126, 0.343]***#	0.211 [0.097, 0.306]***#
	Agri./Cul./Rec.	−0.274 [−0.508, −0.242]***#	−0.320 [−0.411, −0.219]***#	−0.188 [−0.404, −0.143]***#	−0.228 [−0.323, −0.125]***#
	Other	–	–	–	–
Year	2014	−0.361 [−0.654, −0.313]***#	−0.458 [−0.587, −0.326]***#	−0.351 [−0.643, −0.311]***#	−0.454 [−0.582, −0.318]***#
	2015	−0.069 [−0.293, −0.081]***#	−0.143 [−0.234, −0.049]***#	−0.039 [−0.259, −0.043]**#	−0.116 [−0.199, −0.021]***#
	2016	0.025 [−0.223, 0.062]*	−0.081 [−0.188, 0.021]***	0.053 [−0.190, 0.096]***	−0.057 [−0.162, 0.047]***
	2017	0.117 [−0.095, 0.113]***	0.044 [−0.039, 0.135]***	0.128 [−0.083, 0.119]***	0.054 [−0.030, 0.144]***
	2018	0.024 [−0.169, 0.023]*	−0.031 [−0.108, 0.052]**	0.031 [−0.159, 0.026]**	−0.027 [−0.103, 0.054]*

(*continued*)

Table 3. Continued.

Parameter	Category	Model 1 GNB Beta [2.5th, 97.5th]	Model 2 PiG Beta [2.5th, 97.5th]	Model 3 GNB w./PCs Beta [2.5th, 97.5th]	Model 4 PiG w./PCs Beta [2.5th, 97.5th]
	2019	0.097 [−0.101, 0.088]***	0.044 [−0.043, 0.120]***	0.096 [−0.099, 0.090]***	0.043 [−0.041, 0.120]***
	2020	–		–	–
PCs (% of var)	1 (3.86)			−1.034 [−1.618, −0.398]***#	−1.094 [−1.575, −0.589]***#
	2 (3.27)			0.765 [0.158, 1.338]***#	0.754 [0.320, 1.251]***#
	3 (2.47)			1.332 [0.645, 2.024]***#	1.432 [0.909, 1.980]***#
	4 (2.21)			0.735 [0.088, 1.500]***#	0.806 [0.230, 1.440]***#
	5 (1.98)			−0.775 [−1.438, −0.095]***#	−0.628 [−1.205, −0.052]***#
	6 (1.87)			0.105 [−0.597, 0.707]	0.092 [−0.485, 0.690]
	7 (1.76)			−0.519 [−1.082, 0.284]***	−0.636 [−1.187, −0.030]***#
	8 (1.63)			0.396 [−0.376, 1.299]***	0.441 [−0.226, 1.134]***
	16 (1.06)			−1.084 [−2.406, −0.263]***#	−0.932 [−1.727, −0.128]***#
	17 (1.02)			0.953 [0.226, 1.930]***#	1.073 [0.356, 1.815]***#
	21 (0.09)			1.291 [0.304, 2.441]***#	1.456 [0.602, 2.295]***#
	22 (0.08)			−1.353 [−2.591, −0.367]***#	−1.267 [−2.212, −0.375]***#
	28 (0.06)			−0.700 [−2.129, 0.324]***	−0.619 [−1.702, 0.320]***
	29 (0.06)			−0.707 [−2.078, 0.582]***	−0.760 [−1.907, 0.303]***
	31 (0.05)			−0.543 [−1.703, 0.645]***	−0.503 [−1.628, 0.583]***
	32 (0.05)			−1.227 [−2.628, −0.199]***#	−1.330 [−2.447, −0.169]***#
σ			−0.477 [−0.588, −0.380]***#		−0.509 [−0.627, −0.424]***#
φ		0.610 [0.210, 0.828]***#		0.574 [0.183, 0.699]***#	
Q		0.104 [−0.554, 0.327]***		0.081 [−0.616, 0.243]**	
−2LL/Deviance		411,192	396,360	409,723	394,911
AIC		411,250	396,416	409,813	394,999
BIC		411,519	396,675	410,230	395,406
N		76,579	76,579	76,579	76,579

Note: Coefficients were estimated from the whole sample, while BCIs were obtained from 500 bootstrapping replicates with size of 5,000; # indicates the 95% BCIs do not include zero; the model-based p-values are indicated by $^{+}p < .1$, $^{*}p < .05$, $^{**}p < .01$, and $^{***}p < .001$.

−0.053] and −.225 [−0.347, −0.103] in models 1 and 2, respectively. While the significance was not robust against bootstrapping when the content-related PCs were controlled in models 3 and 4. Thus, caution was needed for the interpretation, because the choice of model and the method of bootstrapping sampling may have increased the chance of false negatives even when the same model specification is utilized.

Table 4 presents the estimated coefficients obtained for the Chengdu data. As noted above, the number of working days in Chengdu data possibly had heaped values due to government regulations. Indeed, the proportion of value 5 occupied 24.43% of the cases, while the proportions of values 10 and 20 were only 3.09 and .15%, respectively. Therefore, a group of models that could address both issues of the potential inflation of value 5 and heavy tail was utilized. Models 1 and 2 implemented inflated versions of NB and GNB models using the control variables and measures of the content-agnostic features, and the selected PCs were added to Models 3 and 4 along with the controls and measures of the content-agnostic features.

The number of working days showed a clear declining trend over the years. For instance, according to Model 1, the number of expected working days for a response in 2014 was 1.37 [exp(.862)−1 = 1.37] times higher than that in 2019 or 2020, while the ratio dropped to 1.25 [exp(.231) = 1.26] in 2018. For the type of queried agency, only posts related to public transportation were responded to faster than others—roughly 28% of the time [1 − exp(−.326) = .28], and the effect was robust against bootstrapping evaluation. In addition, there was clear evidence that the dependent variable had a heaping point for value 5, indicated by the significant intercept of the inflation part across all models. Although effects, such as the length of the post and the level of positive sentiment showed a similar pattern to those reported in Table 3, none were considered robust, because of their failure to pass the bootstrapping evaluations.

To further explore the possible meaning of PCs, Table 5 reports the five keywords for the top five topics according to their contributions to the first five PCs. It appeared that PC1 and PC2 did not have a clear meaning for the Luzhou data, because the top topic was about how a question is raised (topic 1300) or how news is reported (topic 501). PC1 and PC2 for the Chengdu data were more related to substantive issues, such as local traffic (topic 1334) for PC1 and property management (topic 1773) for PC2. PC3 and PC4 for the Luzhou data shared three topics, namely, topic 1763 (household regulation), 540 (urban developer), and 186 (housing provident fund), although their contributions differed. Besides the shared topics, PC3 was more about property transactions (topic 1012) and local traffic (topic 1334), while PC4 was related to property management (topic 1773) and news reports (topic 501). Apart from

Table 4. Estimated coefficients for the number of working days to receive a response for the Chengdu data.

		Model 1 Inflated NB	Model 2 Inflated GNB	Model 3 Inflated NB w/PCs	Model 4 Inflated GNB w/PCs
Parameter	Category	Beta [2.5th, 97.5th]	Beta [2.5th, 97.5th]	Beta [2.5th, 97.5th]	Beta [2.5th, 97.5th]
Intercept		0.744 [0.289, 1.218]***#	0.769 [0.248, 1.346]***#	0.843 [0.340, 1.323]***#	0.862 [0.283, 1.402]***#
LogT		0.081 [0.006, 0.162]***#	0.090 [−0.006, 0.173]***	0.074 [−0.001, 0.151]***	0.081 [−0.014, 0.151]***
LogW		0.032 [−0.109, 0.182]	−0.010 [−0.174, 0.179]	0.019 [−0.137, 0.177]	−0.020 [−0.182, 0.185]
Positive Sent		−0.232 [−1.425, 0.898]	−0.533 [−1.639, 0.804]$^{+}$	−0.255 [−1.394, 0.853]	−0.576 [−1.613, 0.683]*
PropS		−0.012 [−0.184, 0.137]	0.033 [−0.168, 0.207]	0.010 [−0.163, 0.160]	0.056 [−0.151, 0.221]$^{+}$
Err		−0.074 [−0.197, 0.052]*	−0.098 [−0.236, 0.059]***	−0.069 [−0.192, 0.064]*	−0.092 [−0.227, 0.082]**
LAWS		0.022 [−0.073, 0.115]	0.045 [−0.077, 0.174]*	0.009 [−0.086, 0.107]	0.030 [−0.090, 0.155]
Med Sent.		−0.171 [−0.358, 0.043]***	−0.175 [−0.412, 0.060]***	−0.150 [−0.345, 0.072]**	−0.143 [−0.345, 0.101]***
MAD Sent.		0.143 [−0.188, 0.431]$^{+}$	0.183 [−0.240, 0.568]*	0.152 [−0.163, 0.456]$^{+}$	0.189 [−0.200, 0.556]*
Med JSD		−0.247 [−0.719, 0.160]*	−0.285 [−0.794, 0.210]**	−0.263 [−0.732, 0.155]*	−0.287 [−0.829, 0.208]**
MAD JSD		0.579 [−0.713, 1.740]*	0.707 [−0.736, 1.989]**	0.502 [−0.792, 1.663]$^{+}$	0.647 [−0.857, 1.932]*
Year	2014	0.862 [0.734, 0.975]***#	0.936 [0.769, 1.057]***#	0.850 [0.711, 0.954]***#	0.929 [0.754, 1.029]***#
	2015	0.735 [0.609, 0.882]***#	0.738 [0.578, 0.890]***#	0.719 [0.589, 0.864]***#	0.727 [0.566, 0.880]***#
	2016	0.933 [0.826, 1.035]***#	0.929 [0.798, 1.019]***#	0.923 [0.806, 1.019]***#	0.926 [0.779, 1.005]***#
	2017	0.760 [0.602, 0.895]***#	0.792 [0.610, 0.959]***#	0.751 [0.596, 0.885]***#	0.793 [0.608, 0.950]***#
	2018	0.231 [0.110, 0.341]***#	0.239 [0.102, 0.377]***#	0.227 [0.107, 0.345]***#	0.236 [0.102, 0.390]***#
	2019/2020			–	–
Queried agency	Public transp.	−0.326 [−0.472, −0.194]***#	−0.328 [−0.478, −0.144]***#	−0.330 [−0.491, −0.173]***#	−0.345 [−0.500, −0.124]***#
	Public safety	0.015 [−0.141, 0.171]	−0.054 [−0.218, 0.130]	0.026 [−0.127, 0.185]	−0.050 [−0.210, 0.133]
	Urban const.	−0.028 [−0.189, 0.127]	−0.044 [−0.216, 0.130]$^{+}$	−0.064 [−0.223, 0.095]*	−0.071 [−0.245, 0.114]*
	Welfare	−0.136 [−0.307, 0.034]***	−0.149 [−0.330, 0.061]***	−0.120 [−0.285, 0.050]***	−0.143 [−0.330, 0.071]***
	Planning	−0.096 [−0.237, 0.045]***	−0.095 [−0.268, 0.089]***	−0.089 [−0.232, 0.055]**	−0.083 [−0.244, 0.111]**
	Adm./Law	−0.043 [−0.193, 0.115]	−0.054 [−0.220, 0.132]*	−0.050 [−0.195, 0.091]$^{+}$	−0.064 [−0.235, 0.117]*
	Agri./Cul./Rec.			–	–
Level of agency	Prefectural city	−0.103 [−0.235, 0.053]**	−0.075 [−0.201, 0.089]*	−0.100 [−0.230, 0.055]**	−0.066 [−0.192, 0.113]*
	Prefecture	0.062 [−0.040, 0.182]**	0.045 [−0.082, 0.174]*	0.054 [−0.046, 0.182]*	0.035 [−0.092, 0.167]
	City			–	–
Inflation @5		−1.593 [−1.701, −1.501]***#	−1.603 [−1.717, −1.517]***#	−1.594 [−1.702, −1.502]***#	−1.605 [−1.727, −1.524]***#

PCs (% of var)	1 (7.96)			−0.253 [−0.777, 0.323]*	0.027 [−0.613, 0.941]
	2 (4.22)			−0.053 [−0.626, 0.512]	0.055 [−0.564, 0.693]
	3 (3.41)			−0.053 [−0.591, 0.658]	0.026 [−0.675, 0.809]
	4 (3.02)			−0.757 [−1.446, −0.044]***#	−0.574 [−1.365, 0.310]***
	5 (2.57)			−0.484 [−1.516, 0.349]**	−0.269 [−1.184, 0.668]+
	6 (2.27)			−0.730 [−1.550, 0.105]***	−0.845 [−1.819, 0.102]***
	7 (2.16)			−0.081 [−0.835, 0.742]	−0.211 [−1.081, 0.621]
	10 (1.61)			−0.559 [−1.419, 0.416]*	−0.757 [−1.627, 0.311]***
	11 (1.13)			−0.351 [−1.412, 0.870]	−0.323 [−1.388, 1.059]
	12 (1.07)			0.046 [−1.005, 1.225]	−0.055 [−1.323, 1.344]
	14 (0.93)			1.124 [−0.079, 2.663]***	1.338 [−0.137, 2.754]***
	15 (0.86)			0.396 [−0.679, 1.562]	0.598 [−0.821, 1.967]*
	16 (0.83)			0.189 [−0.901, 1.253]	0.097 [−1.317, 1.417]
	17 (0.80)			0.565 [−0.548, 2.069]+	0.712 [−0.906, 2.460]*
	18 (0.78)			−0.160 [−1.547, 1.176]	−0.265 [−1.876, 1.447]
α		0.368 [0.313, 0.411]***#		0.365 [0.307, 0.404]***#	
φ			0.049 [0.006, 0.083]***#		0.040 [0.002, 0.058]***#
Q			−1.057 [−2.129, −0.744]***#		−1.156 [−2.664, −0.931]***#
−2 LL		73,319	73,174	73,235	73,069
AIC		73,371	73,228	73,317	73,153
BIC		73,568	73,432	73,626	73,470
N		14,109	14,109	14,109	14,109

Note: Coefficients were estimated from the whole sample, while BCIs were obtained from 500 bootstrapping replicates with size of 5,000; # indicates the 95% BCIs do not include zero; the model-based *p*-values are indicated by $^{+}p < .1$, $^{*}p < .05$, $^{**}p < .01$, and $^{***}p < .001$.

Table 5. Top five topics and their five keywords for the first five PCs by location.

		Top 5 keywords (prob.)										
		Luzhou					Chengdu					
PCs	Topic # (%)	1	2	3	4	5	Topic # (%)	1	2	3	4	5
1	1300 (88.68)	回答 (0.33)	提问 (0.14)	请问 (0.06)	关注 (0.04)	问答 (0.03)	1334 (95.07)	交通 (0.10)	路段 (0.07)	车辆 (0.07)	通行 (0.06)	路口 (0.05)
	501 (1.93)	罕见 (0.13)	记者 (0.11)	近日 (0.07)	泸州 (0.06)	宜宾 (0.04)	1773 (0.93)	物业 (0.30)	业主 (0.28)	小区 (0.12)	住宅 (0.02)	开发商 (0.02)
	1334 (1.07)	交通 (0.10)	路段 (0.07)	车辆 (0.07)	通行 (0.06)	路口 (0.05)	414 (0.60)	大道 (0.20)	大街 (0.09)	路口 (0.05)	高架 (0.02)	交叉口 (0.02)
	1773 (1.03)	物业 (0.30)	业主 (0.28)	小区 (0.12)	住宅 (0.02)	开发商 (0.02)	1296 (0.44)	小区 (0.49)	保安 (0.07)	花园 (0.06)	居民 (0.06)	住户 (0.04)
	1763 (0.92)	户口 (0.17)	户籍 (0.11)	落户 (0.11)	登记 (0.04)	迁移 (0.03)	532 (0.32)	高峰 (0.18)	出行 (0.12)	时段 (0.12)	拥堵 (0.10)	高峰期 (0.04)
2	501 (84.72)	罕见 (0.13)	记者 (0.11)	近日 (0.07)	泸州 (0.06)	宜宾 (0.04)	57 (50.30)	成都 (0.45)	成都市 (0.08)	四川 (0.04)	天府 (0.02)	南充 (0.02)
	1763 (4.94)	户口 (0.17)	户籍 (0.11)	落户 (0.11)	登记 (0.04)	迁移 (0.03)	1773 (23.32)	物业 (0.30)	业主 (0.28)	小区 (0.12)	住宅 (0.02)	开发商 (0.02)
	1300 (3.02)	回答 (0.33)	提问 (0.14)	请问 (0.06)	关注 (0.04)	问答 (0.03)	1296 (10.57)	小区 (0.49)	保安 (0.07)	花园 (0.06)	居民 (0.06)	住户 (0.04)
	540 (2.22)	开发商 (0.11)	购房 (0.08)	楼盘 (0.08)	买房 (0.07)	购房者 (0.05)	457 (7.54)	公交 (0.32)	公交车 (0.13)	出行 (0.04)	巴士 (0.04)	市民 (0.03)
	1773 (1.67)	物业 (0.30)	业主 (0.28)	小区 (0.12)	住宅 (0.02)	开发商 (0.02)	278 (1.95)	停车 (0.21)	停车场 (0.14)	车位 (0.06)	停车位 (0.06)	停放 (0.05)
3	1763 (29.28)	户口 (0.17)	户籍 (0.11)	落户 (0.11)	登记 (0.04)	迁移 (0.03)	457 (60.73)	公交 (0.32)	公交车 (0.13)	出行 (0.04)	巴士 (0.04)	市民 (0.03)
	540 (25.82)	开发商 (0.11)	购房 (0.08)	楼盘 (0.08)	买房 (0.07)	购房者 (0.05)	57 (24.21)	成都 (0.45)	成都市 (0.08)	四川 (0.04)	天府 (0.02)	南充 (0.02)
	186 (7.55)	公积金 (0.23)	住房 (0.21)	提取 (0.07)	贷款 (0.07)	缴存 (0.06)	1773 (4.89)	物业 (0.30)	业主 (0.28)	小区 (0.12)	住宅 (0.02)	开发商 (0.02)
	1012 (7.14)	办理 (0.59)	手续 (0.25)	相关 (0.02)	登记 (0.02)	申请 (0.01)	616 (1.32)	线路 (0.27)	站点 (0.11)	专线 (0.05)	公里 (0.04)	开通 (0.03)
	1334 (6.91)	交通 (0.10)	路段 (0.07)	车辆 (0.07)	通行 (0.06)	路口 (0.05)	1300 (1.11)	回答 (0.33)	提问 (0.14)	请问 (0.06)	关注 (0.04)	问答 (0.03)
4	1763 (39.82)	户口 (0.17)	户籍 (0.11)	落户 (0.11)	登记 (0.04)	迁移 (0.03)	1773 (46.30)	物业 (0.30)	业主 (0.28)	小区 (0.12)	住宅 (0.02)	开发商 (0.02)
	540 (20.87)	开发商 (0.11)	购房 (0.08)	楼盘 (0.08)	买房 (0.07)	购房者 (0.05)	457 (18.12)	公交 (0.32)	公交车 (0.13)	出行 (0.04)	巴士 (0.04)	市民 (0.03)
	186 (16.36)	公积金 (0.23)	住房 (0.21)	提取 (0.07)	贷款 (0.07)	缴存 (0.06)	57 (14.18)	成都 (0.45)	成都市 (0.08)	四川 (0.04)	天府 (0.02)	南充 (0.02)
	501 (8.78)	罕见 (0.13)	记者 (0.11)	近日 (0.07)	泸州 (0.06)	宜宾 (0.04)	1317 (7.25)	影响 (0.19)	环境 (0.13)	情况 (0.09)	达标 (0.08)	困扰 (0.07)
	1773 (2.04)	物业 (0.30)	业主 (0.28)	小区 (0.12)	住宅 (0.02)	开发商 (0.02)	1296 (2.81)	小区 (0.49)	保安 (0.07)	花园 (0.06)	居民 (0.06)	住户 (0.04)
5	186 (54.42)	公积金 (0.23)	住房 (0.21)	提取 (0.07)	贷款 (0.07)	缴存 (0.06)	1296 (59.67)	小区 (0.49)	保安 (0.07)	花园 (0.06)	居民 (0.06)	住户 (0.04)
	540 (16.00)	开发商 (0.11)	购房 (0.08)	楼盘 (0.08)	买房 (0.07)	购房者 (0.05)	1317 (9.22)	影响 (0.19)	环境 (0.13)	情况 (0.09)	达标 (0.08)	困扰 (0.07)
	1773 (13.77)	物业 (0.30)	业主 (0.28)	小区 (0.12)	住宅 (0.02)	开发商 (0.02)	57 (6.89)	成都 (0.45)	成都市 (0.08)	四川 (0.04)	天府 (0.02)	南充 (0.02)
	1334 (2.63)	交通 (0.10)	路段 (0.07)	车辆 (0.07)	通行 (0.06)	路口 (0.05)	1773 (4.75)	物业 (0.30)	业主 (0.28)	小区 (0.12)	住宅 (0.02)	开发商 (0.02)
	1061 (2.45)	医保 (0.13)	医疗 (0.09)	报销 (0.06)	费用 (0.04)	就医 (0.04)	1300 (3.95)	回答 (0.33)	提问 (0.14)	请问 (0.06)	关注 (0.04)	问答 (0.03)

the shared topics with PC3 and PC4, such as topics 186, 540, and 1334/1773, PC5 gave more weight to health care issues (topic 1061).

Although PC1 to PC5 was highly significant and passed the bootstrapping procedure, as shown in Table 3, the sign of the coefficients for the PCs with overlapping topics was not consistent. For instance, PC1 and PC2 shared four out of the top 5 contributing topics, namely, topics 1300, 501, 1773, and 1763; however, a higher PC1 score was associated with a faster response, while the opposite was true for PC2, according to Models 3 and 4 in Table 3. A similar pattern could be found for PC3 and PC4, which shared three topics with PC5.

Despite one may provide some ad-hoc evaluations based on the keywords of the non-shared topics, such as Topic 1334 (traffic issues) for PC1 and Topic 540 (urban developer) for PC2, it could be misleading to interpret the meaning of the PCs using an incomplete comparison of the topic distributions. Since the number of topics in the LDA word corpus is vast, a comprehensive evaluation of the topic distributions could be challenging, if not entirely impossible, without further simplification of the results. More importantly, the goal of the LDA model available in PaddlePaddle is to identify patterns from a high variety of texts, and the topics could portray subtle or even trivial differences that might not be meaningful or interpretable in social science studies. Therefore, caution was needed in interpreting the findings to avoid arbitrariness introduced by individual researchers.

Discussion and conclusion

The current study evaluated how local governments respond to online public requests in China. The main contributions of this work were 2-fold. On one hand, we utilized computational linguistic tools to separate text features into content-related and content-agnostic measures and evaluated their effects on government responsiveness. Both content-related measures, such as PCs, and content-agnostic measures, such as length and sentiment, were associated with response length. However, due to the complexity of the results obtained from computational tools and the modeling strategy adopted in this study, more work is needed to gauge the possibility of a false negative and the substantial meaning of specific results, such as topics and PCs. The results also supported the findings of previous studies. For example, we found that response patterns varied by location, time, and type of queried agency. Location relates to the unique social economic structure and the resources that local government can invest to channel the public. A single channel may lead to a surge of requests and slower response times, as the Luzhou data showed; however, as demonstrated by the Chengdu data, having more intricate systems that specifically targeted audiences may generally improve the speed of the process but may also create communication silos and potentially reduce public engagement. The reason why the type of queried agency matters was most likely due to the type of request—those related to existing policies were responded to faster than those requiring time to investigate and collect information.

Inconsistent with the previous findings, this study did not find that posts seen by officials as conducive to economic growth were more likely to receive a speedy response compared to social welfare requests. This discrepancy might be due to the data sources used and the operationalization of the concepts. For example, many previous studies utilized posts collected from national forums, and welfare was divided into domains, such as the environment, health care, etc., according to the topics categorized by researchers (e.g., Su and Meng 2016). However, the posts in this study were retrieved from local forums, and we did not categorize them by the pre-trained topic sets due to their large size. Instead, they were grouped according to the queried agency. Posts submitted to agencies related to civil affairs, health commissions, human resources and social security, housing provident funds, unions, religious affairs, and veterans' affairs, were classified as welfare issues. Thus, it was not surprising that the results were not entirely the same.

On the other hand, this work scrutinized several methodological issues for utilizing online data in social science and sought practical solutions. Although many computational tools are widely available, integrating computational components in the social sciences is still in its early stage, and social scientists face serious methodological challenges in statistical inference and interpretation. Traditional methodologies built on analysis for datasets with finite records might not be suitable for online content that can easily reach millions of observations. Unlike traditional data that contain a subset of the population, online data is arguably more comparable to a census. Running statistical models on census records would detect trivial effects and increase the chance of false positives. Thus, answers to questions, such as how to evaluate the uncertainties of estimates obtained from all records, and how to conduct feature selection and perform statistical inference, remained unavailable. To alleviate the issue of false positives and to quantify the uncertainties of the results obtained from computational tools in statistical analysis, we adopted a bootstrapping procedure to construct empirical confidence intervals for the estimates. Although this strategy may have led to false negatives, it offered more confidence in the robustness of the results.

Another challenge was about interpretation. Nowadays data mining tools, such as LDA models have been widely employed in social science studies. However, the risk of misinterpretation increases when the results obtained from computational toolkits are blindly used. The primary purpose of data mining tools is to find a suitable model through a data-driven search with few assumptions about the data, and little interest is paid to formulating a specific hypothesis to be tested (Weiss and Davison 2010). By contrast, most statistical analyses start with a hypothesize-and-test paradigm under specific assumptions about the data and relationship forms (Hand 1998). Thus, the focus of data mining is either prediction or description. To be as accurate as possible, the number of parameters preserved by a data mining model could be very large (Wu et al. 2014). In contrast, conventional statistical modeling summarizes, explains, or predicts using a limited number of parameters for parsimony (Vandekerckhove, Matzke, and Wagenmakers 2014). Thus, directly utilizing the results obtained from data mining tools, such as topic models, and finding a meaningful interpretation of the resolved topics can be difficult. Overlapping topics might not be interpretable or meaningful in a way that social scientists find familiar. In addition, like exploratory tools, such as PCA, topic models do not offer information for statistical inference. Although coherence measures can be used to find a plausible number of topics (Röder, Both, and Hinneburg 2015), such practice becomes unfeasible when the data size is massive, the model specification is complex, and the tuning process for the parameters is complicated. Therefore, we conducted further steps, such as compositional analysis and variable selection, to include the results obtained from the exploratory topic models into an explanatory framework. Furthermore, instead of interpreting the LDA topics, we provided a practical way to include them as controls (e.g., selected PCs), similar to what behavioral geneticists do to control population structures. More importantly, by explicitly modeling the heavy tail and inflation to address the distinct distributional features of the dependent variable, we not only reduced the chance of making biased estimates and incorrect inferences but also offered an option for a sensitivity check against specification errors.

Despite the rigorous adoption of analytical tools, several aspects of the current study were limited. First, the posts retrieved from the local forums were subject to selective deletion. All governmental online channels explicitly state that posts that may violate the law or be inappropriate for the forum's purpose are not allowed and will be removed once found. However, it was unclear how such appropriateness is evaluated and enforced. Since the governmental forums were all manually censored and the criteria for accessing appropriateness or sensitiveness varied (King, Pan, and Roberts 2013), it was difficult to know whether a post would be deleted because of its tone (e.g., strong negative sentiment) in a non-transparent process. Second, due to the continuous updating of the models available in the PaddlePaddle, some of the measures (e.g., topics obtained from the LDA models) in the current study might not be reproducible in future versions. In addition, the response patterns might be substantially different in online and offline contexts, and the two forums retrieved in this study might not be representative of local government responsiveness in general, as they were not randomly selected and unknown reasons might exist for their non-random survival (e.g., the local leader's personal effort). Thus, caution needs to be taken in generalizing our conclusions in the offline context.

Despite these limitations, the current study represents one of the research attempts to make some progress in utilizing recent advances in text mining to examine classic problems in the social sciences. We hope this work can inspire future discussions on integrating computational results into classical social sciences, and developing analytical strategies that can overcome methodological challenges and facilitate hypothesis testing with computational components.

Notes

1. More information can be found at https://wen.lzep.cn/node/all.html.
2. Detailed information can be found at http://12345.chengdu.gov.cn/openWorkList.
3. To ensure numerical stability, we removed a few cases with the dependent variable higher than 200 (three for the Luzhou and seven for the Chengdu data). The cutoff threshold of 200 was chosen by experimentation starting from the highest and so forth until the results were numerically stable.
4. A selectivity analysis of all presented models using the top 10 PCs was also conducted, and the results are available upon request.
5. The results for the NB models and topics estimated from other corpuses are available upon request.

ORCID

Tianji Cai http://orcid.org/0000-0002-8962-2660

References

Askira-Gelman, Irit, and AnthonyL. Barletta. 2008. "A "Quick and Dirty" Website Data Quality Indicator." In Proceedings of the 2nd ACM Workshop on Information Credibility on the Web, 43–46. doi:10.1145/1458527.1458538.

Bartels, Larry M. 2006. "Is the Water Rising? Reflections on Inequality and American Democracy." *Political Science & Politics* 39 (1):39–42. doi:10.1017/S1049096506060057.

Benjamin, Walter. 2012. *The Task of the Translator*. Vol. 9. University of Chicago Press.

Berger, Jonah, and Katherine L. Milkman. 2012. "What Makes Online Content Viral?" *Journal of Marketing Research* 49 (2):192–205. doi:10.1509/jmr.10.0353.

Biber, Douglas, Bethany Gray, and Kornwipa Poonpon. 2011. "Should We Use Characteristics of Conversation to Measure Grammatical Complexity in L2

Writing Development?" *Tesol Quarterly* 45 (1):5–35. doi:10.5054/tq.2011.244483.

Binder, Michael, Matthew Childers, and Natalie Johnson. 2015. "Campaigns and the Mitigation of Framing Effects on Voting Behavior: A Natural and Field Experiment." *Political Behavior* 37 (3):703–722. doi:10.1007/s11109-014-9292-2.

Blei, David M., Y. Ng Andrew, and Michael I. Jordan. 2003. "Latent Dirichlet Allocation." *The Journal of Machine Learning Research* 3:993–1022.

Broockman, David E. 2013. "Black Politicians Are More Intrinsically Motivated to Advance Blacks' Interests: A Field Experiment Manipulating Political Incentives." *American Journal of Political Science* 57 (3):521–536. doi:10.1111/ajps.12018.

Butler, Daniel M., and David E. Broockman. 2011. "Do Politicians Racially Discriminate against Constituents? A Field Experiment on State Legislators." *American Journal of Political Science* 55 (3):463–477. doi:10.1111/j.1540-5907.2011.00515.x.

Cai, Tianji, Yiwei Xia, and Yisu Zhou. 2021. "Generalized Inflated Discrete Models: A Strategy to Work with Multimodal Discrete Distributions." *Sociological Methods & Research* 50 (1):365–400. doi:10.1177/0049124118782535.

Chen, Jidong, Jennifer Pan, and Yiqing Xu. 2016. "Sources of Authoritarian Responsiveness: A Field Experiment in China." *American Journal of Political Science* 60 (2):383–400. doi:10.1111/ajps.12207.

Chen, Xi. 2012. *Social Protest and Contentious Authoritarianism in China*. Cambridge University Press.

Chengdu Yearbook Society. 2019. *Chengdu Yearbook 2019*. Chengdu Yearbook Society.

Crossley, Scott A., Jerry Greenfield, and Danielle S. McNamara. 2008. "Assessing Text Readability Using Cognitively Based Indices." *Tesol Quarterly* 42 (3):475–493. doi:10.1002/j.1545-7249.2008.tb00142.x.

Crossley, Scott A., Kristopher Kyle, and Danielle S. McNamara. 2016. "The Tool for the Automatic Analysis of Text Cohesion (TAACO): Automatic Assessment of Local, Global, and Text Cohesion." *Behavior Research Methods* 48 (4):1227–1237. doi:10.3758/s13428-015-0651-7.

Crosslin, David R., Gerard Tromp, Amber Burt, Daniel S. Kim, Shefali S. Verma, Anastasia M. Lucas, Yuki Bradford, Dana C. Crawford, Sebastian M. Armasu, John A. Heit, et al. 2014. "Controlling for Population Structure and Genotyping Platform Bias in the EMERGE Multi-Institutional Biobank Linked to Electronic Health Records." *Frontiers in Genetics* 5 (Sep):352. doi:10.3389/fgene.2014.00352.

Dilip, R. C., T. Lucas, A. Saarangan, Sagara Sumathipala, and Chinthaka Premachandra. 2018. "Making Online Content Viral through Text Analysis." In 2018 IEEE International Conference on Consumer Electronics (ICCE), 1–6. doi:10.1109/ICCE.2018.8326164.

Distelhorst, Greg, and Yue Hou. 2014. "Ingroup Bias in Official Behavior: A National Field Experiment in China." *Quarterly Journal of Political Science* 9 (2):203–230. doi:10.1561/100.00013110.

Esarey, Ashley. 2015. "Winning Hearts and Minds? Cadres as Microbloggers in China." *Journal of Current Chinese Affairs* 44 (2):69–103. doi:10.1177/186810261504400204.

Gandhi, Jennifer. 2008. *Political Institutions under Dictatorship*. Cambridge University Press.
Gilens, Martin. 2005. "Inequality and Democratic Responsiveness." *Public Opinion Quarterly* 69 (5):778–796. doi:10.1093/poq/nfi058.
Givón, Talmy. 1995. *Functionalism and Grammar*. John Benjamins Publishing.
Government Online Project. 2000. "Assembly for Promoting Government Online of 100 Cities." Beijing.
Grohs, Stephan, Christian Adam, and Christoph Knill. 2016. "Are Some Citizens More Equal than Others? Evidence from a Field Experiment." *Public Administration Review* 76 (1):155–164. doi:10.1111/puar.12439.
Guo, Liang, Scott A. Crossley, and Danielle S. McNamara. 2013. "Predicting Human Judgments of Essay Quality in Both Integrated and Independent Second Language Writing Samples: A Comparison Study." *Assessing Writing* 18 (3):218–238. doi:10.1016/j.asw.2013.05.002.
Gupta, Ramesh C., and S. H. Ong. 2004. "A New Generalization of the Negative Binomial Distribution." *Computational Statistics & Data Analysis* 45 (2): 287–300. doi:10.1016/S0167-9473(02)00301-8.
Gupta, Ramesh C., and S. H. Ong. 2005. "Analysis of Long-Tailed Count Data by Poisson Mixtures." *Communications in Statistics-Theory and Methods* 34 (3): 557–573. doi:10.1081/STA-200052144.
Hand, David J. 1998. "Data Mining: Statistics and More?" *The American Statistician* 52 (2):112–118. doi:10.1080/00031305.1998.10480549.
High, Robin. 2017. "Unconventional Statistical Models with the NLMIXED Procedure." *MWSUG* 2017:1–14.
Jain, Romi. 2017. "China's Economic Development Policies, Challenges and Strategies, 1978–Present: An Overview." *Indian Journal of Asian Affairs* 30 (1/ 2):65–84.
Jia, Lianrui. 2019. "What Public and Whose Opinion? A Study of Chinese Online Public Opinion Analysis." *Communication and the Public* 4 (1):21–34. doi:10.1177/2057047319829584.
Jiang, Di, Yuanfeng Song, Rongzhong Lian, Siqi Bao, Jinhua Peng, Huang He, Hua Wu, Chen Zhang, and Lei Chen. 2018. "Familia: A Configurable Topic Modeling Framework for Industrial Text Engineering." *arXiv Preprint* arXiv: 1808.03733.
Jiang, Junyan, Tianguang Meng, and Qing Zhang. 2019. "From Internet to Social Safety Net: The Policy Consequences of Online Participation in China." *Governance* 32 (3):531–546. doi:10.1111/gove.12391.
Kaplan, Robert M., David A. Chambers, and Russell E. Glasgow. 2014. "Big Data and Large Sample Size: A Cautionary Note on the Potential for Bias." *Clinical and Translational Science* 7 (4):342–346. doi:10.1111/cts.12178.
King, Gary, Jennifer Pan, and Margaret E. Roberts. 2013. "How Censorship in China Allows Government Criticism but Silences Collective Expression." *American Political Science Review* 107 (2):326–343. doi:10.1017/S0003055413000014.
Koplenig, Alexander. 2017. "A Data-Driven Method to Identify (Correlated) Changes in Chronological Corpora." *Journal of Quantitative Linguistics* 24 (4): 289–318. doi:10.1080/09296174.2017.1311447.
Kyle, Kristopher. 2016. "Measuring Syntactic Development in L2 Writing: Fine Grained Indices of Syntactic Complexity and Usage-Based Indices of Syntactic Sophistication." PhD Diss., Georgia State University.

Kyle, Kristopher, and Scott A. Crossley. 2015. "Automatically Assessing Lexical Sophistication: Indices, Tools, Findings, and Application." *TESOL Quarterly* 49 (4):757–786. doi:10.1002/tesq.194.

Lin, Jianhua. 1991. "Divergence Measures Based on the Shannon Entropy." *IEEE Transactions on Information Theory* 37 (1):145–51. doi:10.1109/18.61115.

Loyd, Daisy. 2013. "Obtaining Consent from Young People with Autism to Participate in Research." *British Journal of Learning Disabilities* 41 (2):133–140. doi:10.1111/j.1468-3156.2012.00734.x.

Ma, Yanjun, Dianhai Yu, Tian Wu, and Haifeng Wang. 2019. "PaddlePaddle: An Open-Source Deep Learning Platform from Industrial Practice." *Frontiers in Data Computing* 1 (1):105–115.

Meng, Tianguang, P. Yang, and Z. Su. 2015. "Public Opinion and Local Fiscal Decision Making in Authoritarian China: Based on Survey Experiment to Local Government." *Journal of Public Management* 12 (3):57–68.

Nazari, N., and M. A. Mahdavi. 2019. "A Survey on Automatic Text Summarization." *Journal of AI and Data Mining* 7 (1):121–135.

Norris, John M., and Lourdes Ortega. 2009. "Towards an Organic Approach to Investigating CAF in Instructed SLA: The Case of Complexity." *Applied Linguistics* 30 (4):555–578. doi:10.1093/applin/amp044.

Reilly, James. 2012. *Strong Society, Smart State: The Rise of Public Opinion in China's Japan Policy*. New York, NY: Columbia University Press.

Röder, Michael, Andreas Both, and Alexander Hinneburg. 2015. "Exploring the Space of Topic Coherence Measures." In Proceedings of the Eighth ACM International Conference on Web Search and Data Mining, 399–408. doi:10.1145/2684822.2685324.

Schlaeger, Jesper, and Min Jiang. 2014. "Official Microblogging and Social Management by Local Governments in China." *China Information* 28 (2):189–213. doi:10.1177/0920203X14533901.

Snow, Catherine. 2002. *Reading for Understanding: Toward an R&D Program in Reading Comprehension*. Rand Corporation.

Stasinopoulos, D. Mikis, and Robert A. Rigby. 2007. "Generalized Additive Models for Location Scale and Shape (GAMLSS) in R." *Journal of Statistical Software* 23 (7):1–46. doi:10.18637/jss.v023.i07.

State Council. 2016. *Circular on Further Improving Responses to Public Opinion in Open Government Initiatives*. http://www.gov.cn/xinwen/2016-08/12/content_5099154.htm.

Su, Zheng, and Tianguang Meng. 2016. "Selective Responsiveness: Online Public Demands and Government Responsiveness in Authoritarian China." *Social Science Research* 59:52–67. doi:10.1016/j.ssresearch.2016.04.017.

Sung, Yao-Ting, Kuo-En Chang, and Je-Ming Yang. 2015. "How Effective Are Mobile Devices for Language Learning? A Meta-Analysis." *Educational Research Review* 16:68–84. doi:10.1016/j.edurev.2015.09.001.

Tan, Chenhao, Lillian Lee, and Bo Pang. 2014. "The Effect of Wording on Message Propagation: Topic-and Author-Controlled Natural Experiments on Twitter." *arXiv Preprint* arXiv:1405.1438.

Tian, Hao, Can Gao, Xinyan Xiao, Hao Liu, Bolei He, Hua Wu, Haifeng Wang, and Feng Wu. 2020. "SKEP: Sentiment Knowledge Enhanced Pre-Training for Sentiment Analysis." *arXiv Preprint* arXiv:2005.05635.

Van den Boogaart, K. Gerald, and Raimon Tolosana-Delgado. 2013. *Analyzing Compositional Data with R*. Springer.

Vandekerckhove, Joachim, Dora Matzke, and Eric-Jan Wagenmakers. 2014. "Model Comparison and the Principle of Parsimony." In *The Oxford Handbook of Computational and Mathematical Psychology*, edited by J. R. Busemeyer, Z. Wang, J. T. Townsend, and A. Eidels, 300–319. Oxford University Press.

Wang, Shuna, and Yang Yao. 2007. "Grassroots Democracy and Local Governance: Evidence from Rural China." *World Development* 35 (10): 1635–1649. doi:10.1016/j.worlddev.2006.10.014.

Weijters, Bert, Maggie Geuens, and Niels Schillewaert. 2010. "The Stability of Individual Response Styles." *Psychological Methods* 15 (1):96–110. doi:10.1037/a0018721.

Weiss, GaryM, and BrianD. Davison. 2010. "Data Mining." In *Handbook of Technology Management*. Vol. 2, edited by H. Bidgoli, 542–555. New York, NY: John Wiley and Sons.

Wiley, JoshuaF, Bei Bei, John Trinder, and Rachel Manber. 2014. "Variability as a Predictor: A Bayesian Variability Model for Small Samples and Few Repeated Measures." *arXiv Preprint* ArXiv:1411.2961.

Willmot, Gordon E. 1987. "The Poisson-Inverse Gaussian Distribution as an Alternative to the Negative Binomial." *Scandinavian Actuarial Journal* 1987 (3–4):113–127. doi:10.1080/03461238.1987.10413823.

Wu, Xindong, Xingquan Zhu, Gong Qing Wu, and Ding Wei. 2014. "Data Mining with Big Data." *IEEE Transactions on Knowledge and Data Engineering* 26 (1):97–107. doi:10.1109/TKDE.2013.109.

Xu, Ming. 2022. *Pycorrector: Text Error Correction Tool* (version 0.4.2). https://github.com/shibing624/pycorrector

Zhang, Junhua. 2001. "China's"Government Online" and Attempts to Gain Technical Legitimacy." *ASIEN: The German Journal on Contemporary Asia* 80 (80):93–115.

About the authors

Tianji Cai is a professor of sociology at the University of Macau. His research focuses on quantitative research methods, especially the issues of sampling weights in multilevel and longitudinal models. In addition, he is also interested in integrating genetics and sociology in the studies of social and health behaviors.

Fumin Li is a Ph.D. candidate in the Department of Sociology at the University of Macau. His work focuses specifically on computational social science. His research interests are social computing and text mining.

Jian Zhan is a lecturer in the School of Management at Lanzhou University. His research interests are natural language processing, big data analysis, and computing.

Zhengrong Wang is an associate professor in the School of Publica Administration at Gansu Institute of Political Science and Law. His research focuses on public crisis management.