

Article

Close to Beijing: Geographic Biases in *People's Daily*

Morley J. Weston¹ and Adrian Rauchfleisch^{2,*}¹ Graduate Institute of Geography, National Taiwan University, Taiwan (R.O.C); E-Mail: morley.weston@protonmail.ch² Graduate Institute of Journalism, National Taiwan University, Taiwan (R.O.C); E-Mail: adrian.rauchfleisch@gmail.com

* Corresponding author

Submitted: 31 December 2020 | Accepted: 29 April 2021 | Published: 23 July 2021

Abstract

Inequities in China are reflected within state-run media coverage due to its specific role “guiding public opinion,” and with our study we contribute to the geographic turn in the Chinese context with regard to media and journalism. As a subject of a spatial study, China is unique due to several factors: geographic diversity, authoritarian control, and centralized media. By analyzing text from 53,000 articles published in *People's Daily* (rénmín rìbào, 人民日報) from January 2016 to August 2020, we examine how the amount of news coverage varies by region within China, how topics and sentiments manifest in different places, and how coverage varies with regard to foreign countries. Automated methods were used to detect place names from the articles and geoparse them to specific locations, combining spatial analysis, topic modeling and sentiment analysis to identify geographic biases in news coverage in an authoritarian context. We found remarkably uniform and positive coverage domestically, but substantial differences towards coverage of different foreign countries.

Keywords

authoritarian public; China; Chinese media system; *People's Daily*; propaganda; spatial analysis; news values

Issue

This article is part of the issue “Spaces, Places, and Geographies of Public Spheres” edited by Annie Waldherr (University of Vienna, Austria), Ulrike Klinger (European University Viadrina, Germany) and Barbara Pfetsch (Freie Universität Berlin, Germany / Weizenbaum-Institute for the Networked Society, Germany).

© 2021 by the authors; licensee Cogitatio (Lisbon, Portugal). This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

1. Introduction

While scholars have called for a geographic turn in the Chinese context (Sun, 2010) with regard to media and journalism, few studies really have tried to follow that call, and most existing work has relied on qualitative case studies (e.g., Sun & Chio, 2015; Tong, 2013). In this study, we focus on what Usher (2020) calls “journalists as map-subjects” as we try to visualize and analyze the news coverage of *People's Daily* (rénmín rìbào, 人民日報) through the text's spatial qualities, then connect them loosely with associated news values (Galtung & Ruge, 1965; Harcup & O'Neill, 2017; Oppegaard & Rabby, 2016; Walmsley, 1980). Previous scholarship in China has shown measures of inequality to be heterogeneous based on region and scale of study (e.g., He et al., 2017), and there is a need for nuance in spatial analysis of the country in particular. China's spatial diversity and highly controlled

media lead to an emphasis on some news values that are weighed differently than in the typical Western context (Huan, 2016). Some news values such as proximity are also relevant with regard to geographical bias, and they have often been mentioned in studies with a geographic analysis of news reporting (e.g., Brooker-Gross, 1983; Galander, 2012; Harcup & O'Neill, 2017). Geographic bias (Jones, 2008; Whitney et al., 1989) and media bias in general also exists in Western media systems due to, for example, organizational ownership structures, but we assume that unique biases will exist in the Chinese case, especially in an outlet like *People's Daily*, which represents the official view of the Chinese Communist Party (S. N. Liu & Chang, 2020; Robinson, 1981).

There has always been variation within the Chinese media system, and even in *People's Daily*, there was at times room for subtle criticism (Tan, 1990). However, the situation has clearly worsened in comparison to earlier

leadership periods with Xi Jinping's "repressive-coercive strategies towards critical voices" in combination with a "resurrection of the media propaganda role" (Repnikova, 2017, p. 209). In our study, we are interested in this new leadership period as the role of official media has changed. We will first introduce how spatial analysis has been used in communication science in the past, then focus on the Chinese case and discuss the role of *People's Daily* in China.

To answer our main research question of whether geographic biases can be identified in the coverage of *People's Daily*, we combined different computational approaches. We analyzed all articles published between January 1, 2016, and August 31, 2020. First, a complete sample of the titles, subtitles, and leads (here defined as the first 120 Chinese characters) of all available articles from *People's Daily* was gathered and organized. Second, all place names in China and names of countries around the world were extracted using custom-built geoparsing methods. Third, we performed sentiment analysis on each lead containing a place name to determine the positive or negative tone of the article. Fourth, a topic model will be built using these articles, using different provinces over time as additional variables. Fifth, we used this assembled meta information to draw conclusions about the geographic biases in *People's Daily*. This study thus employs techniques gathered from multiple disciplines to gain insights into the overall trends shaping the spatial variations of *People's Daily's* news coverage.

2. Literature Review

Spatial analysis and the study of the ways that geography influences media and communication is rooted in different strands of communication science research, with some authors focusing on geographic bias (Jones, 2008; Whitney et al., 1989), or geographic and cultural distance as described by Galtung and Ruge (1965) in their news values theory. Even though these existing conceptual references are a good starting point for spatial analysis in communication science, only a few studies have tried to use a geographic information system framework for their analysis. Existing studies that combine journalism studies or communication science with spatial analysis are often published in other academic disciplines, commonly in geography and computer science. Geographer Brooker-Gross (1983), for example, analyzed US television news by referring to Galtung and Ruge's (1965) news values. Of course, there are early exceptions like communication scholar Dominick (1977), who analyzed the geographic bias in TV news in the US.

Spatial analysis of the news media is typically restricted by the geoparsing technologies used, and methods are often re-adapted for each study. While these older studies (e.g., Walmsley, 1980; Whitney et al., 1989) relied on manual content analysis, newer studies such as Johnson's (1997) analysis of geographic and cultural proximity combine more traditional content analy-

sis with computer-assisted techniques. Studies published today rely mainly on entity recognition models that automatically extract names of organizations, persons and locations (Duffy et al., 2020). While entity recognition works well in English and other Western languages, these models usually perform worse with Chinese (Wan et al., 2019). The different existing approaches are used in communication science to analyze digital trace data (see Hoffmann & Heft, 2020, for an overview) but also traditional news coverage (Watanabe, 2018). However, much simpler techniques can get comparable results with texts as formally structured as those in the pages of Chinese official newspapers.

2.1. Geographic Bias in the News Media

Research about locations in the news reporting can be roughly divided into two intertwined strands of research. First, as we have already mentioned above, many studies refer to Galtung and Ruge's (1965) news values as cultural as well as geographic proximity influence the selection of news (e.g., Brooker-Gross, 1983). Such studies describe the impact of news values as potential bias (Brooker-Gross, 1983) or discuss it just as news values without referring to it explicitly as bias (e.g., H. D. Wu, 1998). On the other hand, there are studies that mainly focus on bias in news reporting without explicitly mentioning news values (e.g., Dominick, 1977; Duffy et al., 2020; Jones, 2008; Walmsley, 1980; Whitney et al., 1989). The normative question then is whether such a bias of certain regions or countries being under-reported is something undesirable.

One study (Napoli et al., 2018) of US media found that even in local media outlets, only 17 percent of stories were truly local in nature, and that many communities went completely unreported in their sample. When examining the types of communities that received the most mentions in the news media, they compared the number of mentions against a variety of demographic factors, such as population, median income, ethnic makeup, administrative status, and number of universities, and found that news mentions correlated most strongly with population. Other researchers have tried to identify different variables that explain the variable representation of certain regions in the news coverage. Previous studies have shown mixed results regarding the relevance of GDP when applied to news coverage; H. D. Wu (2000) found that the "clout" variables of population, area and GDP were often predictive of news coverage of different countries, but were often precluded by other factors. A more recent study (Atad, 2017) found area, population and GDP to all be significant variables when examining international news coverage. In journalism surrounding natural disasters, it was found that cultural proximity, geographic closeness and number of deaths in a disaster contributed to the amount of coverage and length of international news stories in American media (Adams, 1986).

2.2. News Values

China is an interesting case with regard to news values. Huan (2016), in an article conducted with private interviews with Chinese journalists, showed how the news values of positivity and eliteness are brought to the fore. It can then be hypothesized that Chinese news, especially in a party-line publication such as *People's Daily*, will uphold news values that benefit China's party elite and deviate from those ideals typically exemplified in newsrooms (Bandurski, 2016; Repnikova, 2017). This means if under- or over-coverage of certain regions can be observed after normalizing the attention with population numbers (Whitney et al., 1989), it might be due to different news values such as eliteness, positivity, negativity, or just technological and economical constraints on reporters.

Thus, we can use spatial data to discover spatial biases and provide insights into the news values used to produce content in *People's Daily*. Based on the work of Jones (2008), we can assume that newsworthy events will occur across China loosely in line with population, and that *People's Daily* will report on different provinces in kind. Where we find greater concentrations or gaps in news coverage, this can be a starting point for more detailed investigation into why a place is covered and to which news values this can be ascribed.

2.3. *People's Daily*

People's Daily has a long history as a leading newspaper in China (Robinson, 1981), as it is the official newspaper of the Central Committee of the Communist Party of China and it thus "represents the Party's orthodoxy" (Ye & Zeldes, 2020, p. 26). Scholars have described *People's Daily* as the "voice of the CCP government" (S. Wu, 2014, p. 974) or "Chinese Communist party" (Robinson, 1981, p. 62) and the "national mouthpiece of the central Party leadership" (Wang et al., 2018, p. 126) that sets the "agenda for the rest of the media" (S. N. Liu & Chang, 2020, p. 347).

The role and form of *People's Daily* has changed during different periods of leadership. Robinson (1981), for example, describes in her analysis of the Chinese media system a shift in the style of *People's Daily* after the cultural revolution towards more critical news away from a purely positive propagandistic style. She compares the style at the beginning of the 80s with the time before by referring back to Schell's (1978) description of *People's Daily* during the cultural revolution who describes the newspaper as without any negative news. After the cultural revolution, the content of *People's Daily* gradually changed as they, for example, stopped printing pictures of leaders, a practice that was even criticized in some articles published in the newspaper (Robinson, 1981). A second major shift could be observed in 1989 during the student protests in Beijing. In his analysis about the role of *People's Daily* during the Tiananmen student protests,

Tan (1990) describes how the newspaper changed for a very short time from an official mouthpiece to a newspaper with open reporting about the protests in Beijing.

The contemporary role of *People's Daily* and the media in general was clearly outlined by Xi Jinping in 2016, when his new media policy approach was presented in *People's Daily*. Bandurski (2016) analyzes this shift and concludes that every aspect with regard to media has now to "revolve around the central priority of advancing the Party's agenda." While Repnikova (2017) is slightly more optimistic in her analysis, stating that there is still some room for critical journalism in specific cases, she also sees a clear shift away from the Hu Jintao leadership period towards a stronger emphasis on "fusion of repression-propaganda strategies" (p. 210).

Studies in communication science focusing on *People's Daily* usually conduct quantitative content analysis or qualitative discourse analysis to focus on one specific issue. The Chinese government's views on many issues have been studied using the paper, such as climate change (Pan et al., 2021), the representation of people with disabilities (Ye & Zeldes, 2020) or queer sexualities (Zhang, 2014), the revival of Confucianism (G. Wu, 1994), citizenship (S. N. Liu & Chang, 2020), democratization (Huang & Chen, 2009) or disease coverage (Yang, 2020).

Almost all of these studies use a diachronic perspective by distinguishing different time periods that represent leadership generations (e.g., S. N. Liu & Chang, 2020). A second strand of literature has compared *People's Daily* to more commercially oriented domestic newspapers (e.g., Wang et al., 2018) or a US media outlet such as the *New York Times* (e.g., Luther & Zhou, 2005; Parsons & Xu, 2001). Of special interest for our own study are A. P. L. Liu's (1974) analysis of the rather negative coverage about the US since the founding days of *People's Daily* as well as Lee's (1981) analysis of the more positive coverage of the US in *People's Daily* in 1979 and 1980 as the diplomatic ties between China and the US normalized. Lee (1981) includes in his content analysis the geographic area that is mentioned with regard to the US in articles. An exception is Wan et al.'s (2019) technical study in computer science that uses all articles published in *People's Daily* in 1998 to test their entity recognition model, which also includes locations.

3. Research Questions

With cosmopolitan coastal regions and an isolated inland, rural and urban areas, and different methods of state control, spatial inequality in China can take many forms. The diverse distribution of population, opportunity and wealth (Morales, 2019) throughout China, paired with the special role of *People's Daily* in the national discourse, make this an interesting area of study. We expect that these inequities will be reflected within *People's Daily's* news coverage and can be explained with geographic proximity (Chang et al., 1987), other demographic factors such as GDP or population (Napoli et al., 2018) or

with more specific editorial decisions that should help to “guide public opinion” (A. Chan, 2007). We are interested in what geographic bias exists while accounting for population (Whitney et al., 1989) and GDP. Therefore, our first research question is:

RQ1: Does the amount of news coverage in a province of China scale with population, GDP or other factors?

The supply of news is not determined by demographics alone, and there are bound to be exceptions to these rules. For example, domestic news can be purely positive to communicate a positive image of China to its citizens (Huan, 2016; Shen & Guo, 2013). On the other hand, negative coverage can be used to target specific contentious regions in China in which minorities challenge the legitimacy of China’s Communist Party (Odgaard & Nielsen, 2014), or it can appear because of the more reader-oriented news value of negativity (Huan, 2016). However, we expect the news to be mostly positive with regard to domestic news and stronger biases in more contentious areas. Using the quantitative analysis as a starting point, we will examine some of the exceptions in detail, and analyze the different topics and sentiments expressed in different parts of the country:

RQ2: What accounts for variations in news coverage across space or time?

An important aspect of news coverage is how local the news stories are, even in the context of a national-level publication. In other contexts, the level of local coverage has been the subject of some discussion, such as

by Dickens et al. (2015). Because we will be able to see the administrative level mentioned in these news stories, we can see how specific these stories are in mentioning real-world locations. We will then examine this granularity and how it relates to the content and locations of the news. We thus are interested whether a bias exists on a more granular level that is usually not considered in typical studies focusing on geographic bias (Jones, 2008; Whitney et al., 1989):

RQ3: How specific are the places mentioned in the news media, and does this specificity vary across China?

In our last research question we focus on the international coverage of *People’s Daily*. Shifts in diplomatic relations have had an especially strong impact in the past on the coverage of foreign countries such as the US and are usually evident in the articles published in *People’s Daily*, such as those by Lee (1981) or A. P. L. Liu (1974):

RQ4: What influences the amount and content of news coverage of the US and other foreign countries?

4. Data and Methods

For our study, we gathered data from 56,226 articles published in *People’s Daily* from January 1, 2016, to August 31, 2020 (see Figure 1). We decided to focus on this time frame as the beginning coincides with Xi Jinping’s most recent changes in media policy (Repnikova, 2017). Additionally, the source of our data, the China National Knowledge Infrastructure Database

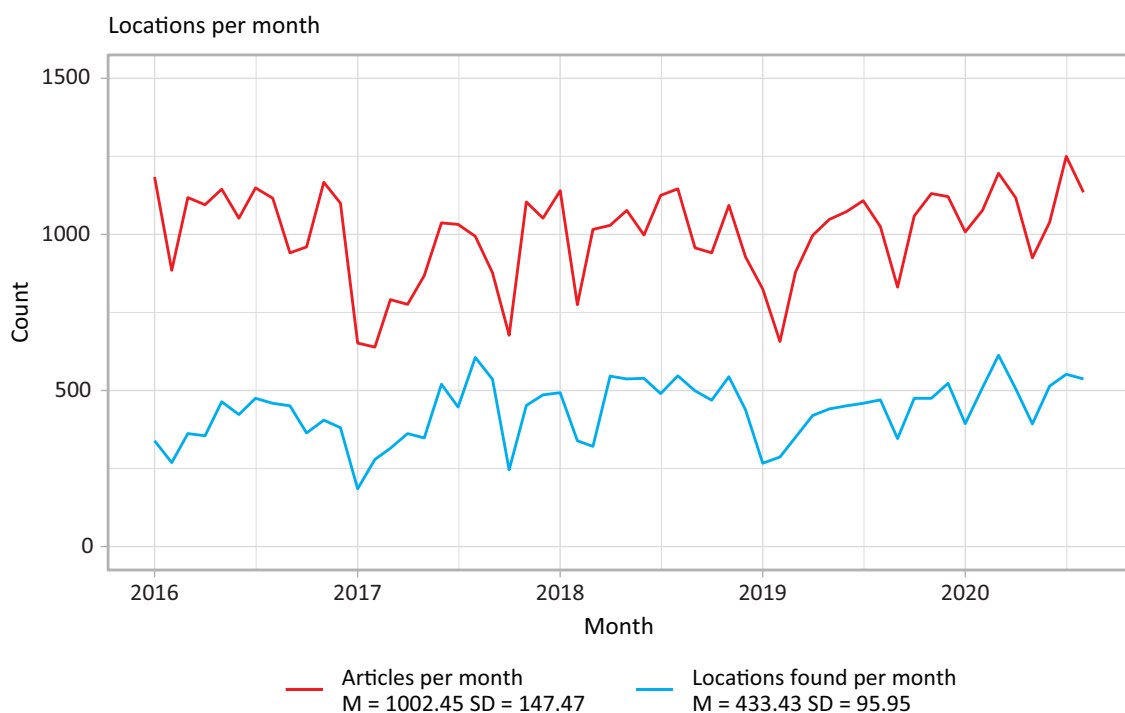


Figure 1. Total articles and recorded locations found per month from January 1, 2016, to August 31, 2020.

(CNKI, 中国知网) has an incomplete dataset for articles before this point in time. The title, author, newspaper section, and first 120 characters were freely available from CNKI's China Core Newspapers Full-text Database, and there was a mean of 1,002 articles gathered per month, with a minimum of 639 and a maximum of 1,250. As *People's Daily* is a printed daily newspaper, the number of articles should be largely consistent, with yearly dips during major holidays. To test whether the first 120 characters were sufficient for this study, we manually compared the geographic names from the leads and the full text of the article and found that the lead was sufficient; of 20 manually inspected samples, 6 contained place names not mentioned in the lead, and none was the primary focus of an article.

To geoparse text, we first considered an entity recognition model (e.g., Sui et al., 2019) that identifies locations as well as organization and personal names. While this approach is better than list-based approaches if person and organization names should be additionally identified (Wan et al., 2019), we decided to use a gazetteer-based approach, as these news stories have a regular structure and a clear hierarchy of place names. China is divided into four administrative levels: the provincial level (省级行政区), prefectural level (地级行政区), county level (县级行政区), and township level (乡级行政区), although names for different regions vary by location and level of autonomy (State Council of the People's Republic of China, 2020), and most smaller locations are accompanied by contextual information in text.

Spatial data from the top two levels of Chinese administrative regions were extracted from a shapefile available from the United Nations Office for the Coordination of Humanitarian Affairs (United Nations Office for the Coordination of Humanitarian Affairs, 2020). In total, 50,563 names corresponding to 46,912 locations were recorded in an SQLite database at all administrative levels. Each article was then scanned for each of these place names using Python, leaving a list of potential place name candidates. The next steps involved narrowing down the list in successive iterations using contextual data, until only the most valid places remained.

We created a coding interface in Python to validate our approach. Against a manually-coded random sample of articles ($N = 127$), this method achieves 96 percent accuracy. While simple, this method performs better than other geoparsing algorithms for similar tasks. Another study (Gritta et al., 2018) of English-language geoparsers tested a number of methods, with results ranging between 66 and 81 percent precision. Our method's superior result can be attributed to several context-specific restrictions. First, *People's Daily* has very predictable, structured content, and almost always uses official place names, which was the main reason to opt for a list-based method (cf. Wan et al., 2019). Second, with only the title, subtitle, and first 128 characters, there is less ambiguity over which identified place names are the most appropriate as the subject of the story.

Some articles were filtered out after the geoparsing process. First, *People's Daily* contains many obituaries and other biographical articles, which contain many locations not relevant to this research in the form of birthplaces or places of education. To eliminate this, any article that contained a reference to the years 1800 to 1980 are identified and eliminated. From the original 56,226 scraped articles, 55,182 were found to be about current events. Second, some articles simply comprise lists of locations, such as those announcing "civilized cities" awards. Because these have little value for this research, any article that contains more than three locations in the lead is filtered out, and we thus reduce our sample to "journalistic" texts. Another 771 articles were filtered out using this method, leaving 54,411 valid articles.

These 54,411 gathered articles contain a total of 24,312 valid place names within Mainland China. 4,110 articles contained more than one location. A plurality of place names recorded were at the prefectural level, with 3,506 of those being references to Beijing City prefecture, the only prefecture in the Beijing City Provincial-level city (see Figure 2). The same basic method was used to search for mentions of different countries around the world, using a gazetteer built from the public domain data set Natural Earth. A total of 16,319 mentions of different countries were found, excluding China.

A variety of sentiment analysis methods were tested, and the most performant was Stanford NLP Group's Stanza natural language processing package in Python which yielded a significant and acceptable correlation of .41 during manual validation ($n = 100$). For details surrounding testing and verification, see the Supplementary File Appendix 1. We also created a structural topic model for the textual data, using those articles that contained place names. This was accomplished using the structural topic model package in R, and captured the top 30 topics across the data set, which we built to contextualize our findings with regard to the third research question. The topic model was validated with a word intrusion test using R's oolong package (C.-h. Chan & Sältzer, 2020) and achieved an acceptable accuracy of 80%, as all topics reported specifically in this article could be successfully identified in word intrusion test.

5. Results

5.1. RQ1: News Coverage by Province

The most basic measure of geographic representation in the news media is the number of news articles mentioning a place. As an absolute measure, this is remarkably uniform; most provincial-level entities received between 1,000 and 2,000 mentions in *People's Daily* over the course of this study (see Figure 3). We compared these numbers to the scaled GDP per capita and scaled population of each province using a Bayesian negative binomial regression model ($N = 30$, R^2 Bayes = .52) and found that if Beijing is excluded, media attention can be

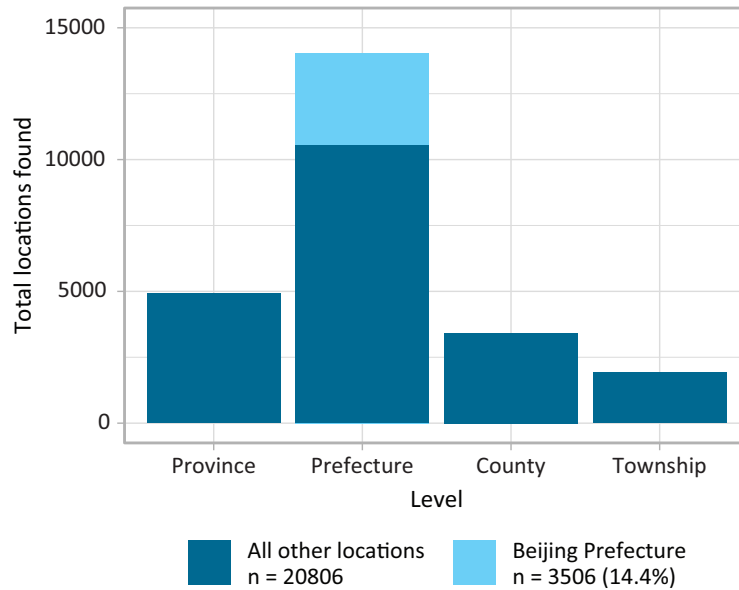


Figure 2. Locations recorded by administrative level in all articles. Notes: n = 24,312. Prefectures were the most commonly mentioned administrative level. Beijing City, the only prefecture in the Provincial-level entity of the same name, was the most commonly mentioned entity.

predicted by GDP per capita (*incidence rate ratio* = 1.25, 95% *CI* = 1.11–1.41) and by population (*irr* = 1.20, 95% *CI* = 1.08–1.33), but Beijing is such an outlier that only GDP per capita (*irr* = 1.47, 95% *CI* = 1.30–1.67) is an acceptable predictor whereas population (*irr* = 1.13, 95% *CI* = .99–1.29) is irrelevant if Beijing is included ($N = 31$, R^2 Bayes = .43). Estimating the model and the marginal effects also allowed us to identify outliers (see Figure 3).

Beijing was the largest outlier in terms of coverage (see Figure 4), but much of this coverage was not about Beijing itself, but about national-level politics centered on Beijing. Because Beijing is such an exceptional case, we examined it separately. A total of 3,746 articles were

found to mention Beijing, of which 3,506 were about the city as a whole. 200 articles that were geoparsed to Beijing were randomly selected; 100 which were of the city as a whole, and 100 of which mentioned a specific district or neighborhood in Beijing. Loosely following the classification laid out by Napoli et al. (2018), they were classified into three categories: (1) Articles not really about Beijing, (2) Articles about events taking place in Beijing but not specifically about the city such as government meetings or diplomatic visits, and (3) those which are relevant to locals of Beijing. An example of category 1 would be an article about a medical team from Beijing going to Wuhan, category 2 might be about a meeting of

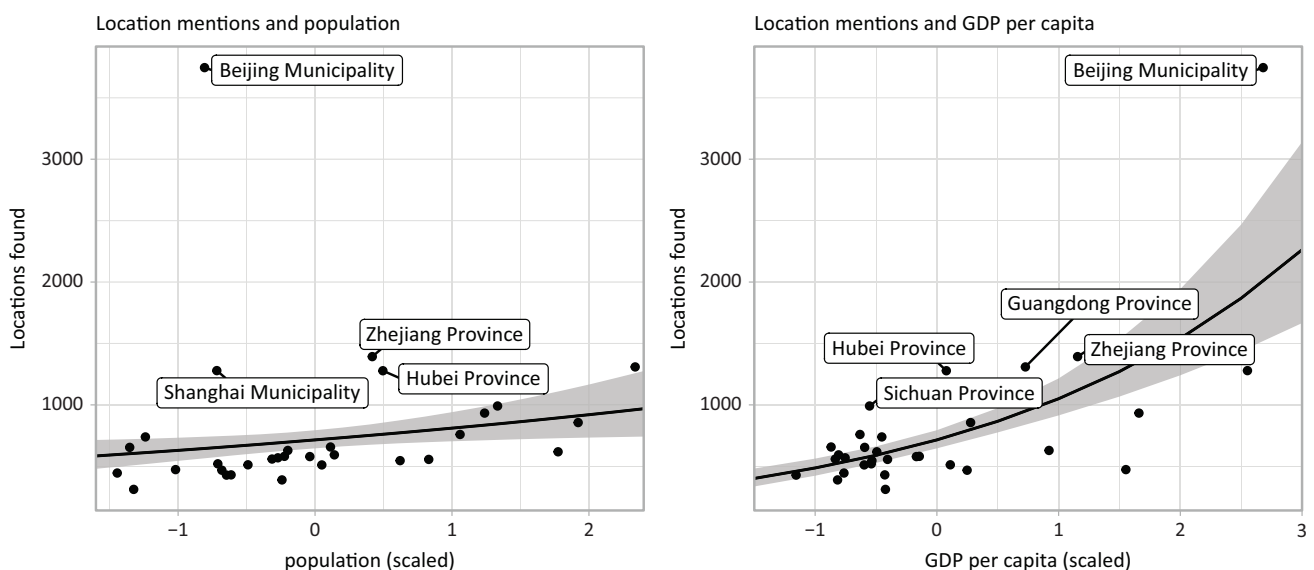


Figure 3. Marginal effects based on the Bayesian negative binomial regression model for population and GDP per capita. Note: Positive outliers are labeled.

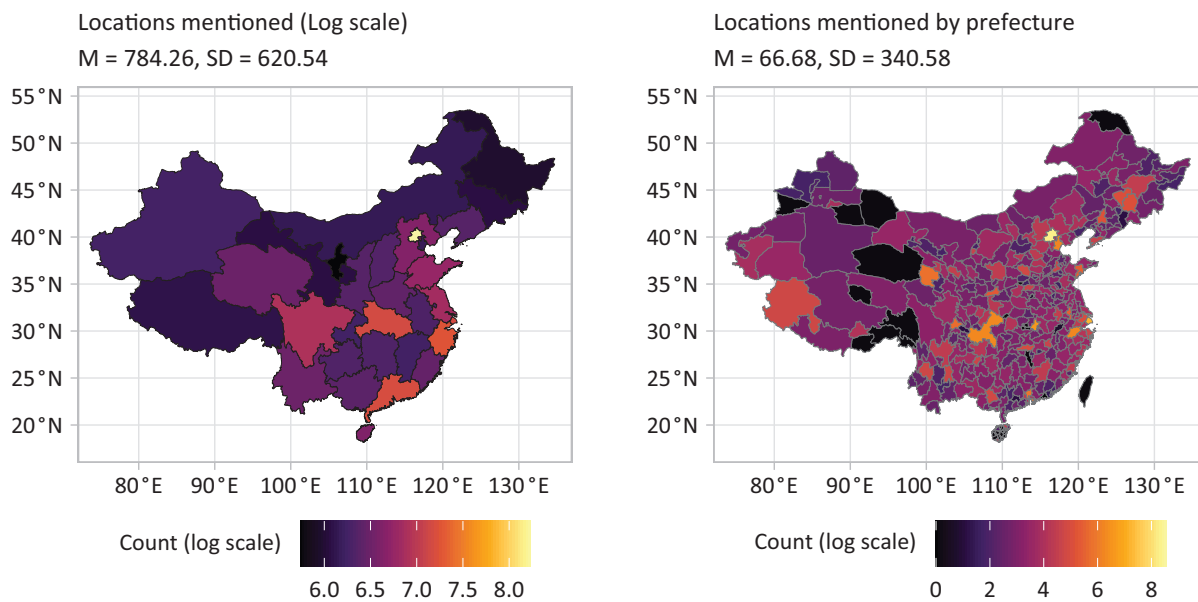


Figure 4. Locations mentioned per province and prefecture. Note: Due to highly variable count data, a natural log scale was used.

the National People’s Congress in Beijing, and category 3 might be about an improvement to a traffic junction in the city. These 200 articles were manually coded to see how locally relevant the articles might be, as well as to what degree citywide and local articles differ. A majority of articles about the city as a whole were about national affairs in Beijing, with 18, 53 and 29 articles in categories 1, 2 and 3, respectively. Those which were about a specific neighborhood were predictably more local, with 10, 36 and 54 percent in the same categories, but there were only 240 of these in the whole data set. If this sample can be extrapolated to the entire data set, weighting for those about the city as a whole and those about individual neighborhoods, there are an estimated 1,146 articles of local Beijing coverage, or 4.7 percent of the entire data set, meaning that Beijing is still over-represented with regard to population.

Two other outliers, Shanghai and Zhejiang, are culturally prominent and wealthy areas of the country. The final outlier, Hubei, was the source of the Coronavirus epidemic, and was genuinely newsworthy during this time period. News mentions of some selected provinces are visualized in Figure 5, as well as reasons for any spikes in news coverage. For example, we can see that the majority of Hubei’s news coverage occurred in the beginning half of 2020, when the Coronavirus was at its peak in Wuhan, but then rapidly declined as time went by (see Figure 5).

5.2. RQ 2: Variations in News Coverage

The initial assumption that the domestic coverage in *People’s Daily* would be overwhelmingly positive were shown to be correct, and this was found to be uniformly true across China. On a scale of -1 (every sentence is neg-

ative) to 1 (every sentence is positive), no provincial average ever approached neutral, and any dips in sentiment were temporary. The mean monthly sentiment of all articles varied very little during this time period ($M = 0.58$, $N = 57$, $SD = 0.06$). Even coverage of Hubei, the origin and epicenter of the Coronavirus pandemic in early 2020, barely dipped below the national average at the height of the pandemic. Sentiment over time in selected provinces is charted in Figure 6, chosen to show the different areas of China, paired with sentiment that rarely diverges from the mean.

5.3. RQ 3: Granularity of Place Name Mentions

We found spatial inequality in the levels of stories from different administrative levels of locations mentioned in different provinces, with the results illustrated in Figure 7. 51 percent of the mentions of Xinjiang were at the provincial level, making it unique among Chinese provincial-level entities.

5.4. RQ 4: International News Coverage and the Role of the US

The same method was used to gather mentions of international locations in the news—articles mentioning different countries were extracted from the news media, and total mentions and sentiment gathered for each. In Figure 8, the total mentions per country are shown.

Using data from the World Bank, the number of mentions per country was compared to population and GDP. Using a linear regression model, a higher population was found to have a statistically significant positive effect on mentions in *People’s Daily* (for full results, see Supplementary File Appendix 3). This contrasts with a

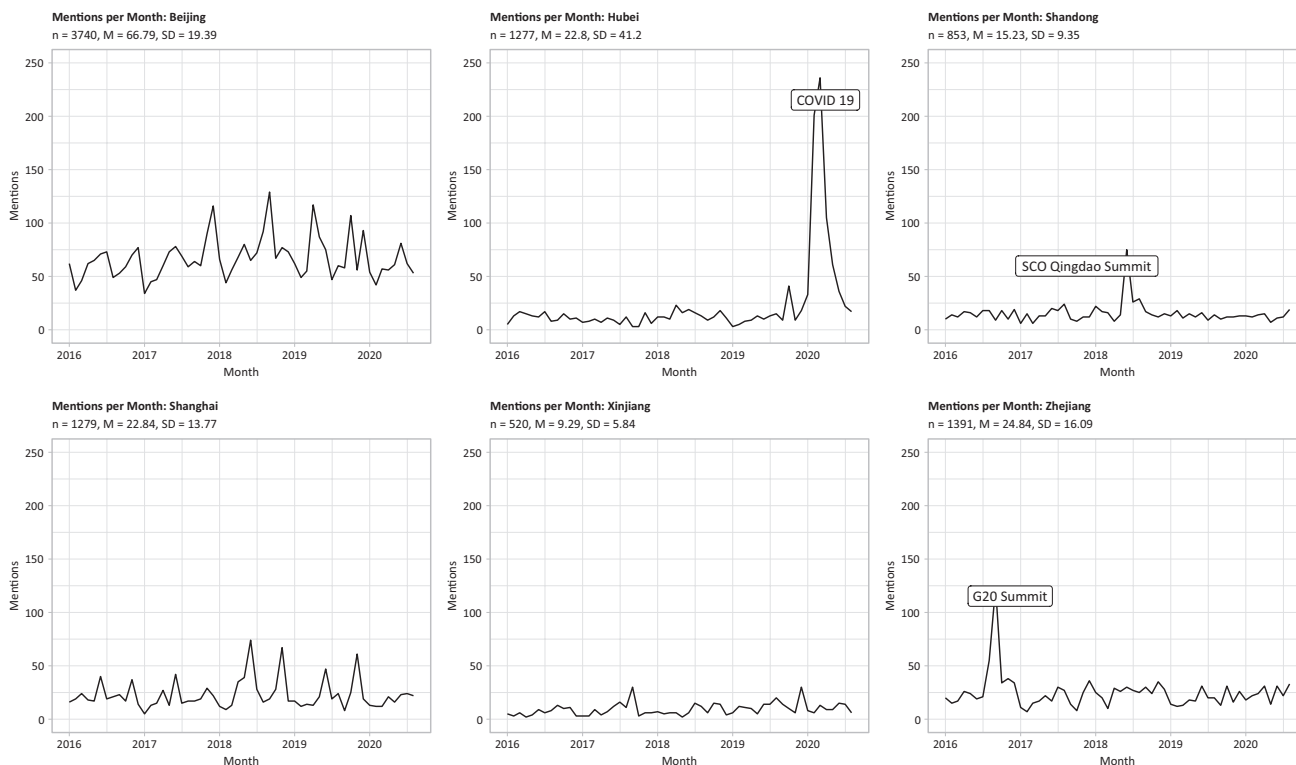


Figure 5. Articles per month for selected locations. Note: Major events are labeled where appropriate.

weak correlation with population in domestic coverage, indicating that in international coverage, *People’s Daily* may function more in line with other news organizations around the world.

The US was by far the most mentioned country in *People’s Daily* except for China itself, and also had some of the most negative coverage. In contrast, most articles mentioning European countries have fairly neutral or positive sentiments. Overall, stories that mentioned international locations scored less positively ($n = 16,319$, $M = .52$, $SD = .52$) than those that mentioned a location within China ($n = 24,312$, $M = .60$, $SD = .46$). The average sentiment per country is shown in Figure 9.

Using a topic model, we can look closer at the difference between the ways that the US and other countries are portrayed in *People’s Daily*. As plotted in Figure 10, we can see 10 selected topics’ relative prevalence in articles mentioning the US and other countries. For a full list of topics, see Supplementary File Appendix 2. From this, we can see that references to Xi Jinping are more common in articles featuring countries other than the US, where the US tends to be clustered as a topic by itself.

6. Discussion

People’s Daily’s coverage of each area of China is remarkably uniform in both amount of coverage and tone. Higher GDP provinces can gain more coverage, and there are some outliers in terms of total coverage. First of all, it can be concluded that with a correlation between the population and the attention each province receives, the

geographic bias usually observed in the media (Jones, 2008; Whitney et al., 1989) is weaker in China. This finding makes *People’s Daily* different from other media organizations studied using similar methods. In the case of *People’s Daily*, it seems that there is a predictable and even amount of coverage for each part of the country.

Regardless of politics or situation on the ground, *People’s Daily’s* domestic news remains uniformly positive in tone. While we could not identify specific trends on a provincial level, we observed a steady and significant overall trend over time (see Supplementary File Appendix 1). In general, our analysis shows that positivity and eliteness are important news values for the *People’s Daily*, and that these standards are rigid across all areas of the country. Our findings are thus in line with the conclusion of Huan’s (2016) qualitative analysis. For example, the lead to one article from 23 June 2020 reads as follows:

Over the past few decades, donkeys have been an important livestock animal in southern Xinjiang, used for travel, hauling and farm work. However, in the past few years, the role of the donkey is slowly changing. In Hotan Prefecture’s Pishan county, Mamat Ulam saw this change. With the improvement of villagers’ lives, the use of donkeys has decreased, and they have decreased in number. Now, the rise of scientific breeding techniques has strengthened the donkey industry, and the lives of villagers have improved. (*People’s Daily*, translated by the authors)

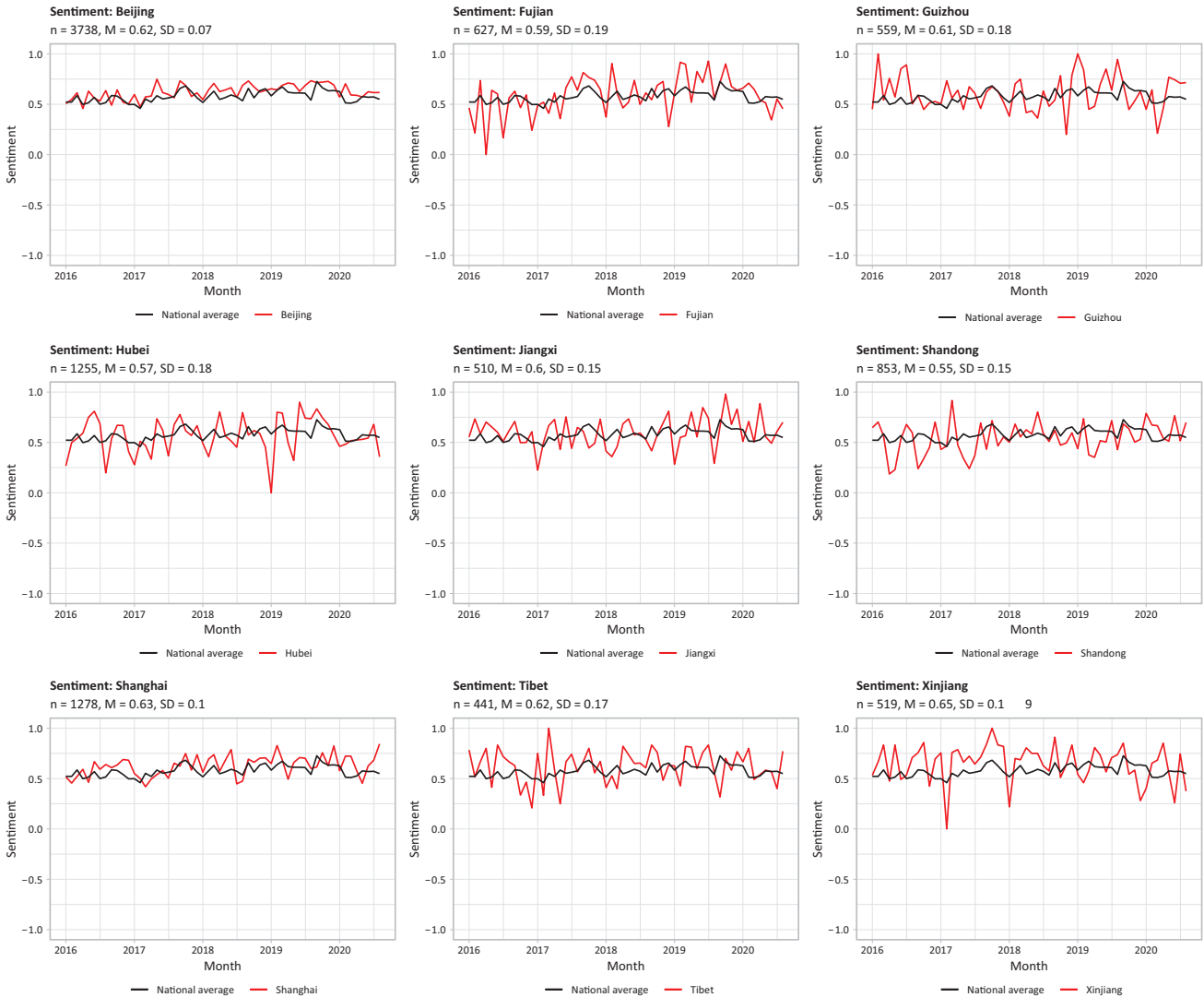


Figure 6. Monthly average sentiment for select provinces. Notes: Over the long term, no news coverage of provinces in China seriously deviated from the mean sentiment values. Some provinces with fewer mentions per month have noisier data and greater variation.

From this short passage, we can see an emphasis on positivity, development and scientific progress, especially towards the improvement of rural citizens’ lives. This also ties into the views of some editorial teams expressed regarding the role of news values in China, where they attempted to hold to the “[t]hree closeness principles (close to the fact, close to daily life, and close to the mass)” (Huan, 2016, p. 4) and the news value positivity (Huan, 2016). However, *People’s Daily’s* focus on positivity does not extend towards international news, especially the US. The coverage is thus different from the time in the late 70s when the diplomatic ties normalized with the US (Lee, 1981).

While there is coverage of negative events in these articles, it is often balanced by positive coverage within a few sentences. For example, one article from July 21 2020 begins:

Since 2004, I have visited Xinjiang in China more than ten times. Xinjiang has beautiful scenery, rich prod-

ucts, and friendly people. I have made many friends there. For a period of time, terrorism, separatism, and extremist forces caused tremendous damage to the stability and development of Xinjiang, posing a serious threat to the lives and property of Xinjiang people. Last year, I was invited to visit Xinjiang again. What happened there? (*People’s Daily*, translated by the authors)

While volume of coverage and sentiment are largely uniform, there are differences in the specificity of news coverage across the country. We can see this reflected in our findings, in which county-level data was more likely to be present in the provinces closest to Beijing. An alternative hypothesis is that because place names in Xinjiang are often transliterated from other languages such as Uighur or Kazakh, they are less likely to be mentioned. However, this pattern is not repeated in Tibet or Inner Mongolia, other regions with their own writing systems. In this sense, Xinjiang is unique among Chinese provincial-level

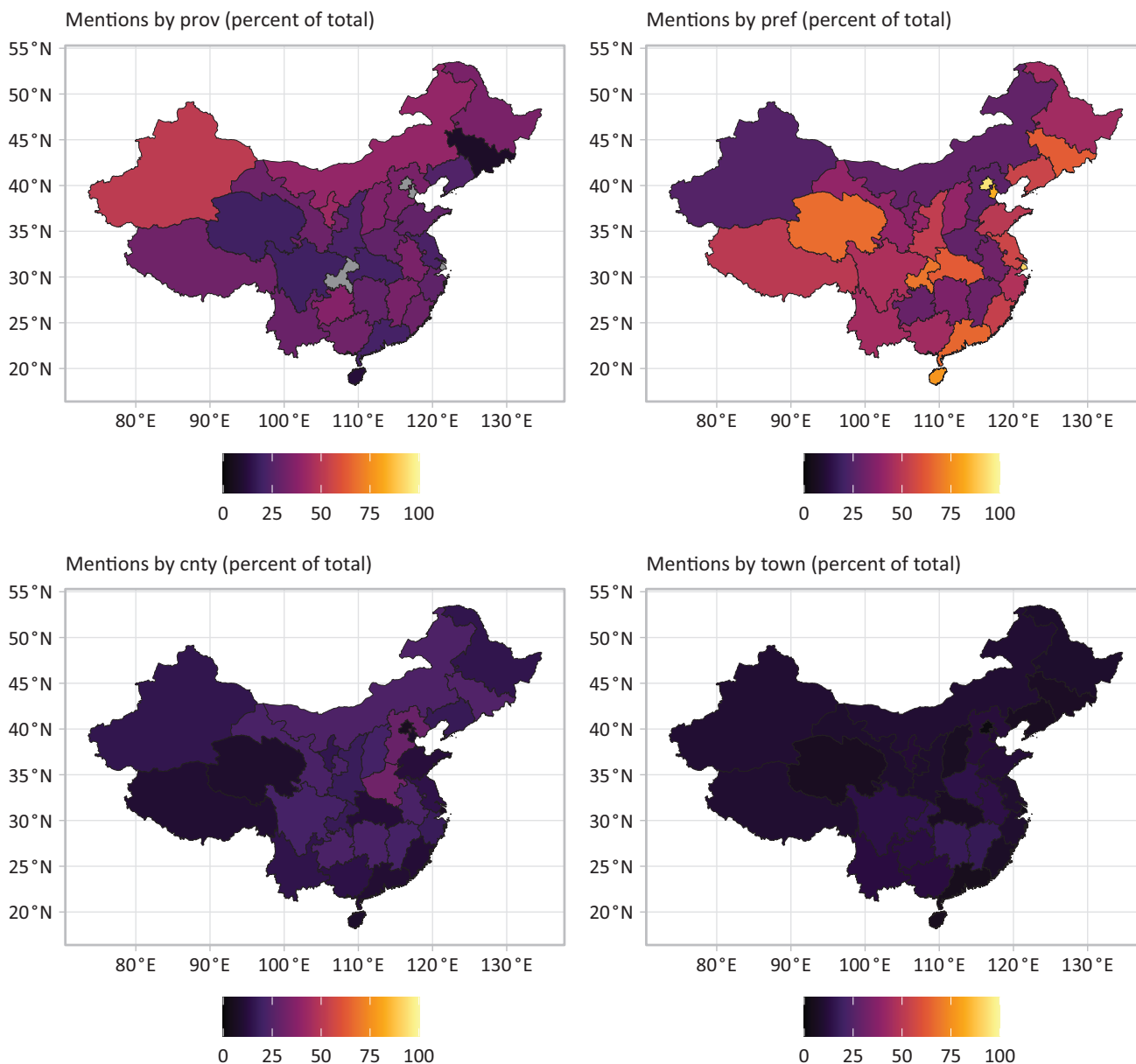


Figure 7. Place name mentions at different administrative levels, aggregated by province. Notes: Provinces are mentioned most often in the news media at different administrative levels. For example, the eastern coastal provinces are most often mentioned at the prefectural level, whereas the northwestern Xinjiang Uighur Autonomous Region is most often mentioned by province.

entities in that it is most often referred to very generally. However, the sentiment values are in line with other parts of the country. This finding could be an indicator of so-called othering as people in Xinjiang are represented as a homogenous group in the coverage (Zhao & Postiglione, 2010). Future qualitative research should analyze in more detail how our findings can be explained.

We also took Beijing as a test case, and were manually able to differentiate between news with a truly local focus and that which took place in the city, but was of a national scope. It can be presumed that this would be true at a lesser extent in other provinces. For example, international events such as the G20 summit in Hangzhou or the Shanghai Cooperation Organization

summit in Qingdao were seen to correlate with spikes in news coverage, but these are not about the cities themselves. This can be seen as a starting point for differentiating truly local and nominally local news coverage using automatic methods.

Our analysis of the specificity of news coverage shows that even if there are no strong geographic variations on the provincial level, there still might be a more granular bias when looking at how much local coverage exists. Many studies focusing on geographic bias (e.g., Jones, 2008; Whitney et al., 1989) or the news value proximity (e.g., Johnson, 1997; H. D. Wu, 1998) have either measured distance to major cities or aggregated the locations to broader regions, but did not consider the

Mentions per country (log scale)
 N = 173, M = 94.33, SD = 215.72

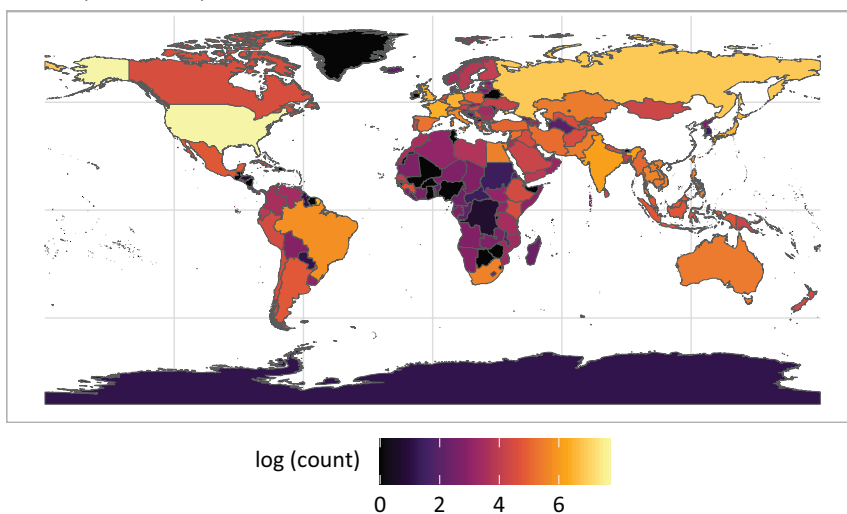


Figure 8. *People's Daily* mentions by country. Notes: Mentions are shown at the country level. The source of this spatial data, Natural Earth, lists all dependent territories separately. Because the count data is highly variable, a natural log scale was used.

Sentiment per country
 N = 134, M = 0.6, SD = 0.18

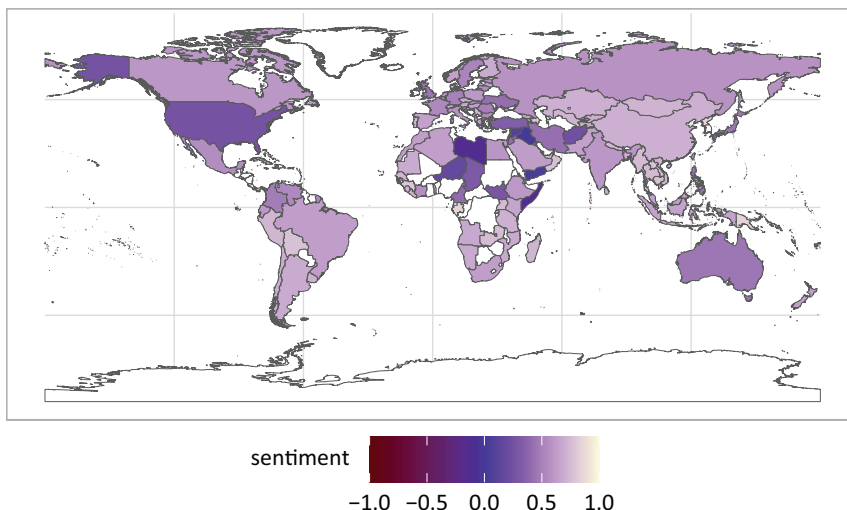


Figure 9. Average sentiment per country. Note: Like in domestic news, no country has an average negative sentiment score, but there is much more variation than at the domestic level.

levels of specificity of place names. Future studies focusing on geographic bias could also include our approach to get a more nuanced picture of which places are covered by news media. However, this is only possible if geoparsers are used that are able to also identify smaller places within a country and assign them a clear hierarchy. Because measures of inequality in China vary by geographic scale (He et al., 2017), this method of analysis could be useful in other contexts.

International coverage in *People's Daily* has much more variation in volume and tone than its domestic counterparts, and in this way is more similar to typical news publications. We can see from this that this newspaper responds to different news values, presumably sub-

ject to different editorial pressures for different types of coverage. The US' exceptional place in news coverage is remarkable, but not unique. H. D. Wu's (2000) study of 38 global newspapers found the US to be the most covered country in the world. The especially negative coverage of the US was likewise expected; it is the subject of many critical editorials in our data set, which can be seen to reflect the CCP's editorial position towards the US (Lee, 1981).

This study is subject to several limitations. First, it only covers one newspaper, so the sample size is limited to how much content is actually available. When dividing the data set into provinces and months, there were often only a few dozen articles per slice, which

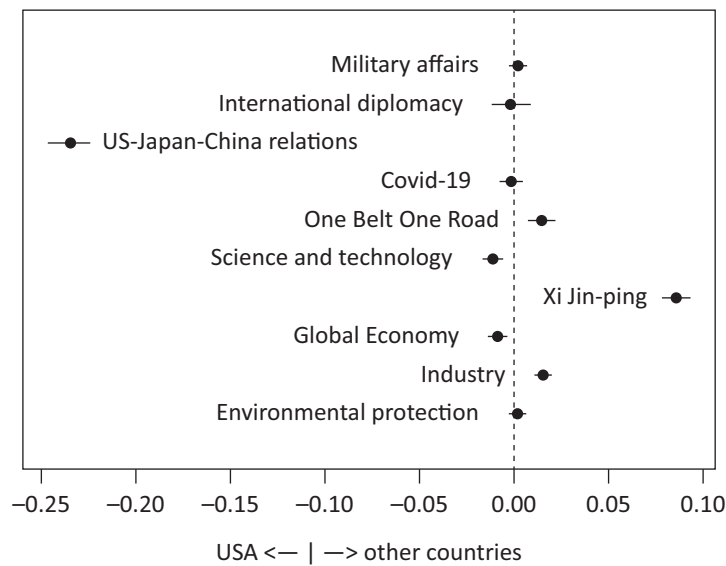


Figure 10. Prevalence difference between articles mentioning the US and other countries. Notes: Topics to the left of the dividing line are more common in articles mentioning the US, while those to the right are more common elsewhere. Lines indicate 95% CI.

introduces a fair amount of “noise” into the data set. Future researchers would be wise to incorporate multiple leading newspapers.

Second, sentiment might not be the most applicable way to judge the intent of Chinese news articles. While it proved to be a useful measure to illustrate perceptions of different countries in the Chinese press, the overall positive tone of domestic coverage meant that the mapped articles were nearly universally positive. Future work could revolve around creating a more applicable typology for Chinese news media, which could better show contrasts between different geographical regions. We believe more qualitative studies of the content will also help to better understand the current editorial strategy of *People’s Daily*.

Acknowledgments

We thank Sie Liang-Yu (謝亮瑜) for her help with the validation of the sentiment analysis and the geoparsing process.

Conflict of Interests

The authors declare no conflict of interests.

Supplementary Material

Supplementary material for this article is available online in the format provided by the author (unedited).

References

Adams, W. C. (1986). Whose lives count? TV coverage of natural disasters. *Journal of Communication*, 36(2),

113–122. <https://doi.org/10.1111/j.1460-2466.1986.tb01429.x>

Atad, E. (2017). Global newsworthiness and reversed domestication: A new theoretical approach in the age of transnational journalism. *Journalism Practice*, 11(6), 760–776. <https://doi.org/10.1080/17512786.2016.1194223>

Bandurski, D. (2016). *Mirror, mirror on the wall*. China Media Project. <https://chinamediaproject.org/2016/02/22/mirror-mirror-on-the-wall>

Brooker-Gross, S. R. (1983). Spatial aspects of newsworthiness. *Geografiska Annaler: Series B, Human Geography*, 65(1), 1–9. <https://doi.org/10.1080/04353684.1983.11879484>

Chan, A. (2007). Guiding public opinion through social agenda-setting: China’s media policy since the 1990s. *Journal of Contemporary China*, 16(53), 547–559. <https://doi.org/10.1080/10670560701562267>

Chan, C.-h., & Sältzer, M. (2020). oolong: An R package for validating automated content analysis tools. *Journal of Open Source Software*, 5(55), Article 2461. <https://doi.org/10.21105/joss.02461>

Chang, T.-K., Shoemaker, P. J., & Brendlinger, N. (1987). Determinants of international news coverage in the U.S. media. *Communication Research*, 14(4), 396–414. <https://doi.org/10.1177/009365087014004002>

Dickens, L., Couldry, N., & Fotopoulou, A. (2015). News in the community? Investigating emerging inter-local spaces of news production/consumption. *Journalism Studies*, 16(1), 97–114. <https://doi.org/10.1080/1461670X.2014.890339>

Dominick, J. R. (1977). Geographic bias in national TV news. *Journal of Communication*, 27(4), 94–99. <https://doi.org/10.1111/j.1460-2466.1977.113-122>

tb01862.x

- Duffy, K. P., Cipparone, H. C., Johnson, E. S., Rickard, L. N., Beard, K., & Nascimento, F. (2020). Leveraging spatial dimensions of news media content analysis to explore place-based differences in natural resource issues. *Journal of Environmental Studies and Sciences*, 10(3), 303–309. <https://doi.org/10.1007/s13412-020-00595-9>
- Galander, M. M. (2012). News values, cultural proximity and cross-cultural media framing: How western and Muslim media covered Darfur. *Journal of Arab & Muslim Media Research*, 5(2), 113–128. https://doi.org/10.1386/jamr.5.2.113_1
- Galtung, J., & Ruge, M. H. (1965). The structure of foreign news: The presentation of the Congo, Cuba and Cyprus crises in four Norwegian newspapers. *Journal of Peace Research*, 2(1), 64–90. <https://doi.org/10.1177/002234336500200104>
- Gritta, M., Pilehvar, M. T., Limsopatham, N., & Collier, N. (2018). What's missing in geographical parsing? *Language Resources and Evaluation*, 52(2), 603–623. <https://doi.org/10.1007/s10579-017-9385-8>
- Harcup, T., & O'Neill, D. (2017). What is news? News values revisited (again). *Journalism Studies*, 18(12), 1470–1488. <https://doi.org/10.1080/1461670X.2016.1150193>
- He, S., Bayrak, M. M., & Lin, H. (2017). A comparative analysis of multi-scalar regional inequality in China. *Geoforum*, 78, 1–11. <https://doi.org/10.1016/j.geoforum.2016.10.021>
- Hoffmann, M., & Heft, A. (2020). “Here, there and everywhere”: Classifying location information in social media data: Possibilities and limitations. *Communication Methods and Measures*, 14(3), 184–203. <https://doi.org/10.1080/19312458.2019.1708282>
- Huan, C. (2016). Leaders or readers, whom to please? News values in the transition of the Chinese press. *Discourse, Context & Media*, 13, 114–121. <https://doi.org/10.1016/j.dcm.2016.05.005>
- Huang, Y., & Chen, J.-p. (2009). Discursive democratization in Mainland China: A diachronic study of the New Year's editorials in the People's Daily. *Journal of Asian Pacific Communication*, 19(2), 199–217. <https://doi.org/10.1075/japc.19.2.07hua>
- Johnson, M. A. (1997). Predicting news flow from Mexico. *Journalism & Mass Communication Quarterly*, 74(2), 315–330. <https://doi.org/10.1177/107769909707400206>
- Jones, S. (2008). Television news: Geographic and source biases, 1982–2004. *International Journal of Communication*, 2, 223–252.
- Lee, C.-C. (1981). The United States as seen through the People's Daily. *Journal of Communication*, 31(4), 92–101. <https://doi.org/10.1111/j.1460-2466.1981.tb00453.x>
- Liu, A. P. L. (1974). Control of public information and its effects on China's foreign affairs. *Asian Survey*, 14(10), 936–951. <https://doi.org/10.2307/2643367>
- Liu, S. N., & Chang, T.-K. (2020). Continuities and changes of media construction of citizenship rights in China: The case of the People's Daily 1978–2012. *Asian Journal of Communication*, 30(5), 343–362. <https://doi.org/10.1080/01292986.2020.1811739>
- Luther, C. A., & Zhou, X. (2005). Within the boundaries of politics: News framing of Sars in China and the United States. *Journalism & Mass Communication Quarterly*, 82(4), 857–872. <https://doi.org/10.1177/107769900508200407>
- Morales, D. X. (2019). Geographic inequality and ethnic stratification in China, 2000 to 2010. *International Journal of Sociology and Social Policy*, 39(7/8), 535–549. <https://doi.org/10.1108/IJSSP-03-2019-0057>
- Napoli, P. M., Weber, M., McCollough, K., & Wang, Q. (2018). Assessing local journalism: News deserts, journalism divides, and the determinants of the robustness of local news. *News Measures Research Project, DeWitt Wallace Center for Media & Democracy*. https://media.kjzz.org/s3fs-public/Assessing-Local-Journalism_100-Communities.pdf
- Odgaard, L., & Nielsen, T. G. (2014). China's counterinsurgency strategy in Tibet and Xinjiang. *Journal of Contemporary China*, 23(87), 535–555. <https://doi.org/10.1080/10670564.2013.843934>
- Oppegaard, B., & Rabby, M. K. (2016). Proximity: Revealing new mobile meanings of a traditional news concept. *Digital Journalism*, 4(5), 621–638. <https://doi.org/10.1080/21670811.2015.1063075>
- Pan, Y., Opgenhaffen, M., & Van Gorp, B. (2021). China's pathway to climate sustainability: A diachronic framing analysis of People's Daily's coverage of climate change (1995–2018). *Environmental Communication*, 15(2), 189–202. <https://doi.org/10.1080/17524032.2020.1817766>
- Parsons, P., & Xu, X. (2001). News framing of the Chinese embassy bombing by the people's daily and the New York Times. *Asian Journal of Communication*, 11(1), 51–67. <https://doi.org/10.1080/01292980109364792>
- Repnikova, M. (2017). *Media politics in China improvising power under authoritarianism*. Cambridge University Press.
- Robinson, D. C. (1981). Changing functions of mass media in the People's Republic of China. *Journal of Communication*, 31(4), 58–73. <https://doi.org/10.1111/j.1460-2466.1981.tb00451.x>
- Schell, O. (1978). *China: In the People's Republic*. Gollancz.
- Shen, F., & Guo, Z. S. (2013). The last refuge of media persuasion: News use, national pride and political trust in China. *Asian Journal of Communication*, 23(2), 135–151. <https://doi.org/10.1080/01292986.2012.725173>
- State Council of the People's Republic of China. (2020). *Zhōnghuá rénmín gònghéguó xíngzhèng qūhuà* [Administrative divisions of the People's Republic

- of China]. http://www.gov.cn/test/2005-06/15/content_18253.htm
- Sui, D., Chen, Y., Liu, K., Zhao, J., & Liu, S. (2019). Leverage lexical knowledge for Chinese named entity recognition via collaborative graph network. In K. Inui, J. Jiang, V. Ng, & X. Wan (Eds.), *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)* (pp. 3828–3838). Association for Computational Linguistics. <https://doi.org/10.18653/v1/D19-1396>
- Sun, W. (2010). Scaling Chinese media: A geographic turn to future research. *International Journal of Communication, 4*, 537–543.
- Sun, W., & Chio, J. (Eds.). (2015). *Mapping media in China: Region, province, locality*. Routledge.
- Tan, F. (1990). The People's Daily: Politics and popular will-journalistic defiance in China during the Spring of 1989. *Pacific Affairs, 63*(2), 151–169. <https://doi.org/10.2307/2759720>
- Tong, J. (2013). The importance of place. *Journalism Practice, 7*(1), 1–16. <https://doi.org/10.1080/17512786.2012.698921>
- United Nations Office for the Coordination of Humanitarian Affairs. (2020). *China: Subnational administrative boundaries* [Data set]. Humdata. <https://data.humdata.org/dataset/china-administrative-boundaries>
- Usher, N. (2020). News cartography and epistemic authority in the era of big data: Journalists as map-makers, map-users, and map-subjects. *New Media & Society, 22*(2), 247–263. <https://doi.org/10.1177/1461444819856909>
- Walmsley, D. J. (1980). Spatial bias in Australian news reporting. *Australian Geographer, 14*(6), 342–349. <https://doi.org/10.1080/00049188008702782>
- Wan, Z., Xie, J., Zhang, W., & Huang, Z. (2019). BiLSTM-CRF Chinese named entity recognition model with attention mechanism. *Journal of Physics: Conference Series, 1302*(3). <https://doi.org/10.1088/1742-6596/1302/3/032056>
- Wang, H., Sparks, C., & Huang, Y. (2018). Measuring differences in the Chinese press: A study of People's Daily and Southern Metropolitan Daily. *Global Media and China, 3*(3), 125–140. <https://doi.org/10.1177/2059436418806022>
- Watanabe, K. (2018). Newsmap: A semi-supervised approach to geographical news classification. *Digital Journalism, 6*(3), 294–309. <https://doi.org/10.1080/21670811.2017.1293487>
- Whitney, D. C., Fritzler, M., Jones, S., Mazzarella, S., & Rakow, L. (1989). Geographic and source biases in network television news 1982–1984. *Journal of Broadcasting & Electronic Media, 33*(2), 159–174. <https://doi.org/10.1080/08838158909364070>
- Wu, G. (1994). Command Communication: The politics of editorial formulation in the People's Daily. *The China Quarterly, 137*(Mar), 194–211.
- Wu, H. D. (1998). Geographic distance and US newspaper coverage of Canada and Mexico. *International Communication Gazette, 60*(3), 253–263. <https://doi.org/10.1177/0016549298060003004>
- Wu, H. D. (2000). Systemic determinants of international news coverage: A comparison of 38 countries. *Journal of Communication, 50*(2), 110–130. <https://doi.org/10.1111/j.1460-2466.2000.tb02844.x>
- Wu, S. (2014). The revival of Confucianism and the CCP's struggle for cultural leadership: A content analysis of the People's Daily 2000–2009. *Journal of Contemporary China, 23*(89), 971–991. <https://doi.org/10.1080/10670564.2014.882624>
- Yang, Z. (2020). Military metaphors in contemporary Chinese disease coverage: A case study of the People's Daily 1946–2019. *Chinese Journal of Communication, 35*(1), 1–19. <https://doi.org/10.1080/17544750.2020.1818593>
- Ye, W., & Zeldes, G. A. (2020). The representation of people with disabilities in an official newspaper in China: A longitudinal study of the People's Daily from 2003 to 2013. *Journal of Disability Policy Studies, 31*(1), 26–34. <https://doi.org/10.1177/1044207319868783>
- Zhang, Q. (2014). Representation of homoeroticism by the People's Daily since 1949. *Sexuality & Culture, 18*(4), 1010–1024. <https://doi.org/10.1007/s12119-014-9237-2>
- Zhao, Z., & Postiglione, G. A. (2010). Representations of ethnic minorities in China's university media. *Discourse: Studies in the Cultural Politics of Education, 31*(3), 319–334. <https://doi.org/10.1080/01596301003786928>

About the Authors



Morley J. Weston holds an MS in geography from National Taiwan University, and is a PhD Student at the Department of Communication and Media Research (IKMZ) at the University of Zurich. A former Journalist in Myanmar and Taiwan, his research interests include tracking media repertoires and developing novel methods of news content analysis.



Adrian Rauchfleisch (PhD, University of Zurich) is an Assistant Professor at the Graduate Institute of Journalism at the National Taiwan University. In his research, he focuses on the interplay of politics, the internet, and journalism in Asia, Europe, and the US.