The China Quarterly (2025), **262**, 531–547
doi:10.1017/S0305741025000177

RESEARCH REPORT

Surveying Internet Users in China: Comparing Representative Survey Data with Official Statistics

Daniela Stockmann[1] and Ting Luo[2]

[1]Hertie School of Governance, Berlin, Germany, and [2]Department of Public Administration and Policy, University of Birmingham, Edgbaston, UK
Corresponding author: Daniela Stockmann; Email: stockmann@hertie-school.org

Abstract
China is well known for providing official data, but how to treat these data is a longstanding debate among China scholars. This paper advances understandings of how to interpret Chinese official statistics about the internet. Using standards for evaluating surveys in the social sciences, we systematically compare official data from the China Network Information Center (CNNIC), which is under the supervision of China's main regulator of internet policy, with the China Internet Survey 2018 (CIS), which is, to our knowledge, the first nationally representative survey on internet use in China. Using three examples, we illustrate how methodological differences in sampling design and measurement can lead to vastly different conclusions about key indicators of internet use in mainland China, including the percentage of internet users, their regional and urban–rural digital divide, and the percentage of specific social media platforms. We discuss the challenges of survey work on internet use in China and offer recommendations on how to interpret official statistics, especially in light of the limitations researchers face when conducting face-to-face surveys in China.

摘要
中国政府有各种各样的官方数据, 但中国研究的学者对如何解读和使用这些数据一直存在争议。本文旨在探讨如何解读中国官方的互联网统计数据。借助社会科学领域中评估问卷的精典标准, 本文系统地比较了两组数据, 中国互联网络信息中心公布的官方数据和我们自己的问卷调查——中国互联网调查2018的数据。据我们所知, 后者是首个针对中国互联网用户的全国代表性调查问卷。通过三个指标, 本文阐明了抽样设计和测量方法的差异会如何得出截然不同的结论。这三个指标分别是, 互联网用户的比例、地区和城乡间的数字鸿沟, 以及社交媒体平台的使用比例。本文也探讨了在面对面问卷调查受限的大背景下, 在中国开展互联网使用调查所面临的挑战, 并提供了关于如何解读官方统计数据的一些建议。

Keywords: China; digital divide; internet; social media; survey
关键词: 中国; 数字鸿沟; 互联网; 社交媒体; 问卷

The Chinese government has become more transparent over time, and more information about China has become available to researchers, in part owing to the internet.[1] One important data source for researchers constitutes official data, which has become easily accessible digitally. While China is a frontrunner in providing official statistics, their reliability is often questioned owing to the absence of freedom of information laws.[2] Empirical research that tests the reliability of official statistics has

1 Stockmann 2010; Stromseth, Malesky and Gueorguiev 2017.
2 See, e.g., Piotrowski et al. 2009.

© The Author(s), 2025. Published by Cambridge University Press on behalf of SOAS University of London. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

https://doi.org/10.1017/S0305741025000177 Published online by Cambridge University Press

shown that some nationwide estimations are systematically biased. For example, the cadre evaluation system provides incentives for local officials to overreport local GDP growth rates.[3] Building on these findings, we aim to advance understandings about Chinese official statistics by focusing on important indicators for internet development.

Statistics provided by the China Network Information Center (CNNIC), which comes under the supervision of the Cyberspace Administration of China (*wangxinban* 网信办), the main regulator of China's internet policy, constitute a major official data source. Tightening controls on data collection and data transfer, combined with the increasing hesitation of Chinese scholars to collaborate with researchers located outside of China, mean that nationally representative surveys as an independent source of information have been put on hold as of 2022.[4] As a result, observers of the Chinese internet and social networks will increasingly need to rely on CNNIC's official statistics to understand China's digital developments.

How should official statistics about China's internet development be understood? Below, building on our experience with the China Internet Survey (hereafter, CIS), we use criteria developed in the social sciences to highlight methodological differences in drawing data.[5] To our knowledge, the CIS is the first nationally representative survey on internet use in China. We find that differences in sampling and measurement between CNNIC and CIS lead to significantly different conclusions about key aspects of internet development, including internet use, regional digital divide and social media use. We aim to enhance knowledge on how to interpret official statistics, with a broader goal of drawing accurate conclusions about China's internet development.

To lay a foundation for our inquiry, we first explain the challenges involved in survey work on internet use and how we addressed these challenges based on best practices developed in survey research on China. We then illustrate how differences in sampling and measurement lead to vastly differing substantive conclusions about internet use and access to social media platforms. We conclude with recommendations on how to interpret official internet statistics, especially considering the limitations researchers face in conducting face-to-face surveys in China.

Challenges for Surveys on Internet Users in China

Official internet statistics are already playing an increasing role in research on China, as alternative data collection methods have become more challenging over time. Below, we place recent developments within the broader context of survey work introduced to China in the 1980s.[6]

In many developing countries, survey sampling based on officially registered residents creates bias. China's economic reform witnessed massive rural-to-urban migration, estimated at roughly 27 per cent of the population in 2020.[7] Since many migrants do not have *hukou* 户口 (household registration), lists of officially registered citizens exclude a large part of the population.[8] To address this challenge, Pierre Landry and Mingming Shen introduced GPS random sampling, which samples survey respondents based on their location.[9] This technique is now not only adopted in China but also widely used in other developing contexts in Africa and South Asia. Google and Baidu maps have greatly facilitated GPS random sampling since those maps can be used to distinguish between inhabited and uninhabited areas, such as, for example, deserts or mountain regions, thus aiding the random sampling of relevant research locations. The CIS 2018 relied on GPS random sampling, with some adjustments for capturing internet users.

3 Plekhanov 2017; Wallace 2022.
4 Greitens and Truex 2020.
5 Stockmann and Luo 2018.
6 Manion (2010) observed 204 institutions in 2004, 83 in 2005, 58 in 2006, and 186 in 2007.
7 National Bureau of Statistics 2021.
8 Solinger 1999.
9 Landry and Shen 2005.

https://doi.org/10.1017/S0305741025000177 Published online by Cambridge University Press

Table 1. Comparison of Privacy Behaviour Questions in the Draft and Final Versions of CIS 2018

Question: Please recall if you have done the following when using social media	
Draft survey	*Final survey*
1. 更改用户简介的隐私设置来限制别人可以看到的信息 Change my privacy settings to limit the information others can see	1. 更改用户简介的隐私设置来限制别人可以看到的信息 Change my privacy settings to limit the information others can see
2. 不让某些人看到您发布的某些东西 Not let some people see what you publish	2. 不让某些人看到您发布的某些东西 Not let some people see what you publish
3. 将某人从您的社交媒体中删除或者拉黑 Blacklist or delete someone	3. 将某人从您的社交媒体中删除或者拉黑 Blacklist or delete someone
4. 删除或重新编辑您以前发过的内容 Delete or edit what you have posted before	4. 删除或重新编辑您以前发过的内容 Delete or edit what you have posted before
5. 删除别人的评论 Delete someone's comments	5. 删除别人的评论 Delete someone's comments
6. 要求别人删除社交媒体上关于您的信息, 包括图片或视频 Ask others to delete their posts, such as photos or videos of you	
7. 忽略或拒绝别人的好友申请 Ignore or reject others' friend requests	

Another key challenge relates to the design of a survey on sensitive issues. Researchers affiliated with institutions outside of China are by law required to collaborate with institutions situated in China. In order to conduct foreign-affiliated surveys, permission must first be granted by the relevant government authorities. As a result, survey design and questionnaires need to be prepared in anticipation of obtaining approval or adjusting accordingly when the National Bureau of Statistics (NBS) requests changes. Only approved mainland institutions listed on the NBS website are permitted to cooperate with foreign researchers.[10] Usually, Chinese collaborators have long-standing connections to these government units and provide foreign researchers with feedback about what can or cannot be asked before submitting the survey questionnaire for government review. What is considered politically sensitive varies depending on the audience, time and region; interviews in Tibet and Xinjiang are usually not an option.[11] When we conducted the CIS, we were asked to reduce sensitive survey questions twice when our survey was delayed owing to the changing political climate in China. For example, we designed and pre-tested seven questions to measure privacy behaviour. Almost all items focus on behaviour with a negative connotation, i.e. rejecting, deleting, blacklisting, limiting, and we were asked to reduce the number of negative items to avoid arousing suspicion about our intentions in using this set of questions. Table 1 compares the questions used in the draft survey and the final survey. Some of the questions included in the CIS 2018 would likely be considered too sensitive from 2019 onwards.

Preference falsification presents another key challenge. This is when citizens refuse to answer or lie when responding to survey questions they consider to be sensitive. China scholars have developed many useful techniques to reduce bias resulting from preference falsification. For example, a 2004 study of political efficacy in China reveals significant bias when using vignettes,[12] while list experiments reveal response bias concerning questions on political trust.[13] Contrary to commonly

10 Manion 2010.
11 Tsai 2010.
12 King et al. 2004.
13 Tang 2016; Robinson and Tannenberg 2019.

https://doi.org/10.1017/S0305741025000177 Published online by Cambridge University Press

held assumptions, experimental evidence has demonstrated that such response bias was likely influenced by factors other than political fear, at least during the Hu–Wen administration.[14] Alternatively, citizens may perceive questions to be intrusive or they may be concerned that their answer to the question will be considered socially undesirable.[15] In the CIS 2018, we limited response bias by reducing the sensitivity of the questions and conducted extensive qualitative interviews and pre-tests to ask questions in non-sensitive ways.

There are two examples of asking questions in non-sensitive ways in the CIS 2018. We were aware that if we asked respondents directly if they had bypassed the Chinese Firewall, they might tend to underreport this behaviour, as it is not allowed by the state. Therefore, we opted for a non-sensitive way to measure such behaviour and instead asked respondents if they had accessed the following platforms: Twitter, Facebook, Reuters Chinese version, Google, BBC Chinese version, *Financial Times*, WhatsApp, Skype, *New York Times* or LinkedIn. As of 2018, Twitter, Facebook, Reuters China, Google, BBC Chinese version, WhatsApp and *New York Times* were blocked in mainland China, while Reuters Chinese version, Skype and LinkedIn were still accessible. Accessing the blocked platforms or websites would require circumventing the Chinese Firewall.[16] By asking about respondents' access to platforms or websites, we made the question less sensitive and more comfortable for them to answer. To reduce the sensitivity of the question addressing privacy concerns, we inquired about "other people's" concerns about privacy rather than the respondent's own concerns.[17] We asked about the reasons why "other people" would be careful when expressing opinions or having political discussions. Furthermore, in qualitative interviews, we used euphemisms commonly employed by internet users to express concerns about state repression, such as "checking the water meter" (*chashuibiao* 查水表) and delivering parcels (*songkuaidi* 送快递), which refer to house visits by the police. Interviewees were more comfortable discussing state censorship and repression using these terms.

A final challenge relates to the overall response rates of surveys. Response rates in China are usually extraordinarily high compared to other countries. For example, the response rate for the CIS was 67.1 per cent, while the response rate for a comparable public opinion survey in the UK, the British Election Study Post-Election Face-to-Face Survey in 2017, was 46.2 per cent.[18] However, response rates can decline in certain locations. The performance of Chinese officials is appraised according to the cadre-evaluation system, which sets the priorities for their localities, the most important of which are economic growth and social stability.[19] Thus, local government officials may not allow data collection out of fear that sensitive information may become public and trigger social unrest or lead to criticism from higher-level authorities. Such repercussions would negatively impact their performance evaluations and advancement within the administrative hierarchy. The use of iPads to conduct surveys, as compared to printed questionnaires, makes it even harder to gain local officials' trust and permission. Scrolling down a survey on a tablet takes more time, providing local officials with additional opportunities to scrutinize questions in greater detail and overinterpret the sensitivity of some questions – for example, those that ask respondents to assess the government. After detecting this issue in a pre-test, we used paper-based questionnaires for the CIS 2018.

Yet the above methods offer little assistance to researchers when government institutions block cross-border collaborations between researchers when conducting nationally representative surveys. One reason for stopping collaboration with researchers outside of China concerns data transfer. Since the 1990 Four-County study – a collaboration between academics at the University of Michigan and Peking University funded by the American National Science Foundation – it has been common

14 Lei and Lu 2017; Stockmann, Esarey and Zhang 2018.
15 Tourangeau and Yan 2007.
16 For results, see Stockmann and Luo forthcoming.
17 Alpert 1971; Fisher 1993.
18 Fieldhouse et al. 2019.
19 Birney 2014.

https://doi.org/10.1017/S0305741025000177 Published online by Cambridge University Press

practice for survey questionnaires to be coded in China, as these were regarded by the authorities as "state secrets," while coded data was able to cross borders freely.[20] China's Data Security Law, which came into force in September 2021, put a stop to this pragmatic solution by stipulating that the export of "important data" collected in China must follow security management protocols set by state authorities. Furthermore, a "Draft data export security assessment," issued by China's Cyberspace Administration agency in October 2021, required risk self-assessments and state approval for any data being transferred across borders.[21] These tighter restrictions, combined with the shifts in the domestic climate that led to foreigners being depicted as potential spies, have left domestic collaborators hesitant to collaborate with researchers located outside of mainland China.[22]

The decline in opportunities to collaborate with Chinese institutions has led to less primary data being collected. Subsequently, researchers have become increasingly dependent on official statistics when investigating the Chinese internet. CNNIC already constitutes a major data source for journalists as well as academics. A search of the Social Science Citation Index between 1945 and 2019 showed that 385 articles on the Chinese internet cited CNNIC reports. A search on Lexis Nexis, which covers over 35,000 media sources worldwide, indicates that between 2000 and 2021, 7,939 English-language news articles on China cited CNNIC reports.[23] These articles rarely discuss the problems associated with official statistics, such as manipulation of data, especially on politically sensitive items and during politically sensitive times.[24] Below, we systematically compare sampling of CNNIC with CIS data to shed light on potential bias resulting from relying on official data.

CNNIC Data

CNNIC is an administrative agency responsible for the operation, administration and services of the internet. It has carried out regular surveys on internet development in China since 1997, releasing biannual statistical reports. In 2014, the organization was placed under China's Cyberspace Administration agency. For comparison, we use the CNNIC report published in February 2019, as it is based on data collected roughly during the same period as that for the CIS 2018. CNNIC does not mention when its survey was conducted; however, the report specifies that the deadline for the statistical survey data is 31 December 2018. Based on this time frame, the survey was probably conducted close to the end of 2018. CNNIC data were collected using computer-assisted telephone interviews.[25] The target population of interest comprised Chinese permanent residents aged six or above (including minors and those over 66 years of age). Random sampling was carried out based on fixed landline (including home telephones and dormitory telephones) and mobile phone numbers, acknowledging the potential overlap of sampling the same person twice. CNNIC adopted stratified two-stage sampling. To ensure that the sample was sufficiently representative of the population, the Chinese mainland was divided into 31 strata, with each province, municipality and autonomous region as

20 Marshall 1993; Manion 2010.

21 "Shuju chujing anquan pinggu banfa zhengqiu yijian gao" (Draft data export security assessments for public consultations). Cyberspace Administration of China, 2021, http://www.cac.gov.cn/2021-10/29/c_1637102874600858.htm. Accessed 21 January 2022; "Zhonghua renmin gongheguo shuju anquan fa" (Data Security Law of the People's Republic of China). National People's Congress, 2021, http://www.npc.gov.cn/npc/c30834/202106/7c9af12f51334a73b56d7938f99a788a.shtml. Accessed 21 January 2022.

22 "Shou ge quanmin guojia anquan jiaoyu ri pufa xuanchuan huodong xingshi duoyang" (The first national security education day featured diverse forms of legal education and publicity activities). *Guangming ribao*, 15 April 2016, https://epaper.gmw.cn/gmrb/html/2016-04/15/nw.D110000gmrb_20160415_3-01.htm. Accessed 11 February 2025; "Guo'an jiguan pilu neimu! Shi guan neidixuesheng! Jintian, zhege shipin yiding yao kan!" (National security authorities disclose inside information! It's about mainland students! Today, you *must* watch this video!). *Pengpai*, 15 April 2021, https://www.thepaper.cn/newsDetail_forward_12214266. Accessed 25 February 2022.

23 For access to Lexis Nexis news articles, please see https://www.lexisnexis.com/en-us/professional/academic/nexis-uni.page.

24 Wallace 2016; 2022.

25 CNNIC 2019.

https://doi.org/10.1017/S0305741025000177 Published online by Cambridge University Press

an independent stratum. In other words, all 31 provinces, municipalities and autonomous regions were sampled. However, CNNIC does not provide crucial information on response rate, sample size, size of overlaps, the sample size of each stratum and how those factors are taken into account, for example, via weights, when presenting descriptive statistics. It thus remains unknown how CNNIC arrives at its final estimates. Moreover, CNNIC's sampling strategy excludes people who have neither a landline nor mobile phone but who can still access the internet (for example, via internet cafes). According to our estimates based on CIS data, this group make up about 4 per cent of Chinese internet users. In sum, although CNNIC data are collected from all provinces and even the adolescent internet population, it is important to acknowledge that CNNIC statistics are based on a random sample of internet users in China. The sampling design means that the adult internet population is underestimated.

CIS Survey Data

The China Internet Survey (CIS 2018) was carried out from July to September 2018 via face-to-face interviews conducted by interviewers affiliated with a Chinese research institute. The population of interest comprised Chinese permanent residents from 18 to 65 years of age.[26] The survey was based on a multi-stage probability spatial sample designed to be representative of the entire Chinese mainland population. Stratification took into account a combination of geography and the degree of internet penetration at the provincial level. First, 75 township-level units were drawn by probability proportional to size (PPS). Each sampled township was then gridded into half arc-minute cells and matched to the population-density model of WorldPop to account for variability in population density within each township.[27] Two secondary sampling units (SSUs) were drawn in each primary sampling unit (PSU), one from the upper quintile of cell density, the other from the remainder. This design led to a share of urban respondents of about 50 per cent, which is comparable to the 2010 population census data. In the absence of further disaggregated density data, tertiary sampling units (namely, hectares, TSU hereafter) were drawn by simple random sampling (SRS) within each SSU and randomly sorted. Field supervisors were instructed to locate the five nearest households from the centre of the first listed hectare, proceed with interviews and continue down the randomly sorted list of hectares until the target of 20 respondents per SSU (roughly, a village) was reached. Thus, the sample properly captured all residents regardless of registration or migration status and employed the latest remote-sensing data to determine unit measures of size below the PSU.[28] A total of 4,686 eligible households were drawn, yielding a final sample size of 3,144 respondents. The response rate was 67.1 per cent. Survey questionnaires took an average of 44 minutes to complete, with all responses recorded via paper-and-pencil on a printed questionnaire. At the analysis stage, we computed a set of post-stratification weights to calibrate the data to be representative of the working-age Chinese population, based on age, gender and education information, which was derived from the 1% National Population Sample Survey conducted by the National Bureau of Statistics in 2015. Table 2 presents the summary statistics of key variables calculated both with and without survey weights.

Beijing was made a "self-representing PSU stratum," which resulted in a separate survey being conducted in Beijing between December 2017 and April 2018 for the most important questions contained in the CIS. The Beijing survey was based on a multi-stage probability spatial sample designed to be representative of Beijing adult residents. The target population was Chinese citizens aged 18 and above who had resided in the urban core and new surrounding districts of Beijing for at least six months. A total of 5,018 eligible samples were drawn. The response rate was 50 per cent, resulting in a final sample size of 2,522. Survey weights were calculated based on demographic information – age and gender – from the 1% Beijing Population Sample Survey conducted in 2015.

26 Chinese nationals who have lived in the community for at least half a year.
27 WorldPop 2022.
28 Landry and Shen 2005.

https://doi.org/10.1017/S0305741025000177 Published online by Cambridge University Press

Table 2. Summary Statistics

				Without survey weights		With survey weights	
Variables	*N*	*Min*	*Max*	*Mean*	*SD*	*Mean*	*Linearized Std. Err.*
Wealth (possession index)	3,109	0	1	0.32	0.25	0.36	0.01
Urban	3,144	0	1	0.73	0.44	0.74	0.03
Internet user	3,144	0	1	0.52	0.50	0.71	0.01
Rural migrant worker	3,132	0	1	0.24	0.43	0.31	0.02
Education	3,139	1	5	2.81	1.28	3.21	0.00
Male	3,144	0	1	0.49	0.50	0.50	0.00
Age	3,144	18	65	46.31	13.86	40.56	0.00

Notes: Wealth (possession index) was an index created based on household ownership of consumer durables and properties using Multiple Correspondence Analysis. The index included the following 8 items: car made in China, plasma or LCD TV, microwave, iPad or tablet, laptop, smartphone, air conditioner and ownership of at least two properties. Rural migrant worker was a dummy variable that took a value of 1 if the respondent was a migrant worker and 0 otherwise. Migrant worker status was identified by two questions: status of household registration and current occupation. Respondents were considered migrant workers if they had a rural (agricultural) household registration but worked in the non-agricultural sector (excluding government, police bureau and the military). Education was measured on a 5-point ordinal scale indicating whether participants had completed lower than primary school education, primary school education, junior high school education, senior or secondary vocational high school education, or college education and above. The measurements of Urban and Internet user are introduced in detail in later sections of this article.

Comparison of the CIS to Existing Surveys on Internet Users

Prior research on internet users has relied on nationally representative surveys that contained questions about internet use.[29] The CIS, to our knowledge, is the first nationally representative survey primarily aimed at studying internet use in China. In terms of spatial sampling, China's four municipalities (Beijing, Tianjin, Shanghai and Chongqing) formed one self-representing stratum, while the remaining 28 provinces were distributed into 15 strata based on their geographic location (east, west, south, north and central) and the number of IP (version 4) addresses per capita divided into three layers.[30] Within each stratum, the number of PSUs (township-level units) was proportional to the percentage of the stratum population in the total population according to the 2010 census data. We deliberately chose a sample that did not include Beijing in the municipality stratum and sampled Beijing separately (see above), thus overall including two municipalities – Beijing and Chongqing. In doing so, the CIS was designed to provide representative estimates across locations that differed with respect to the spread of the internet, as indicated by using number of IP addresses to design the strata. To arrive at the most accurate estimates, the CIS also used fine-grained population data, including WorldPop data that incorporated the sixth China census, GIS data and satellite images to enumerate population structures and estimate population distributions in every half square minute as well as in every 100 m*100 m grid in China.[31] These were used for drawing SSUs and TSUs.

Unlike the CIS, which was collected in-person with a face-to-face questionnaire, research on the Chinese internet tends to use online surveys, most of which are collected through online panels of internet users accessed by international survey companies.[32] One of the challenges facing these surveys is the self-selection of respondents into these panels. The existing studies often use CNNIC reports for quota sampling based on socio-demographics, such as gender or age, to counter this bias. As a result, the generalizability of findings drawn from online surveys, when compared to CIS survey data, is more limited.

The above explanations of survey design matter for the conclusions drawn about internet use in China. Below, we use three key aspects of internet development to illustrate this point. Given the

29 Truex 2017; Roberts 2018.

30 National Bureau of Statistics 2015.

31 WorldPop 2022.

32 See, e.g., Su, Liu and Zhou 2020; Zhang and Guo 2021; Zhou, Su and Liu 2021; Guo, Lei, and Chen 2021.

https://doi.org/10.1017/S0305741025000177 Published online by Cambridge University Press

importance of CNNIC official statistics in academic research on Chinese internet use, our focus is on comparing the differences in conclusions drawn from CNNIC data with those drawn from CIS data.

Example 1: Understanding Internet Use

The widespread use of smartphones in China has led some experts to suspect that 100 per cent of the Chinese population is already connected to the internet. People may be unaware that they are accessing the internet, which may lead to the potential underreporting of internet use. In the CIS, we asked respondents about their activities on a computer or mobile phone. We coded the following items as internet use: sending/reading emails; reading news; looking for information; online shopping; listening to music online; online gaming; watching TV series/movies online; doing business online; reading E-books; using WeChat, QQ or Weibo; using Taobao, Alipay or online banking; and online stock trading. In addition, we also asked whether the respondent "accessed the internet through their own device or the device of someone else." When comparing responses to these questions, only 25 out of 3,144 respondents did not self-report as internet users but were then defined by us as internet users based on how they used computers or smartphones. This strengthens our confidence that self-reports – the most common measure for internet use – is a valid measure for internet usage in China. Below, we rely on a combined measure, which includes the people who were unaware of their internet use.

In comparison, CNNIC defines internet users as "Chinese residents at the age of six or above who have used the internet in the past six months."[33] The sampling population therefore includes adolescents and is larger than that of the CIS. Additionally, CNNIC's sampling design excludes individuals who access the internet neither via mobile nor wired connection at home, which accounts for 4 per cent of internet users according to CIS estimates (62 out of 1,627). These elements in sampling design significantly affect the final estimates.

Figure 1 shows the percentage of internet users between 2000 and 2023. The black bars represent CNNIC data and the grey bars represent the CIS 2018 data. According to the CIS estimates, 71 per cent of the Chinese population used the internet in 2018. These estimates are significantly higher than CNNIC estimates, which report only 56 per cent for 2018. Since CCNIC incorporates a larger population of interest, we attribute these differences to the exclusion of certain internet users. The CIS data reveal that the Chinese internet population may be significantly larger than the current estimates based on official statistics.

Example 2: Understanding the Digital Divide

Research on the digital divide studies how digital technology reinforces socio-economic divisions within society.[34] One important aspect of the digital divide refers to internet access and use.[35]

Research on China's digital divide in internet use has detected a strong regional divide in terms of internet use between the more developed east coast and the less developed western regions, which reflects trends in economic development.[36] So far, this has led to the assumption that internet penetration rates and economic development (in terms of GDP) are directly related. Wei Song, for example, classifies provinces based on their respective internet development rates and draws the conclusion that the country can be hierarchically divided into three tiers: east, middle and west.[37] Notably, in several reports, CNNIC statistics paint a similar picture: "Owing to the gap in local economic development and internet infrastructure construction, the internet penetration rate

33 CNNIC 2019, 92.
34 Dutton, Rogers and Jun 1987; Norris 2001.
35 Livingstone and Helsper 2007; van Deursen and van Dijk 2014.
36 See, e.g., Oi 1999; Huang 2008; Naughton 2018.
37 Song, Wei 2008.

https://doi.org/10.1017/S0305741025000177 Published online by Cambridge University Press

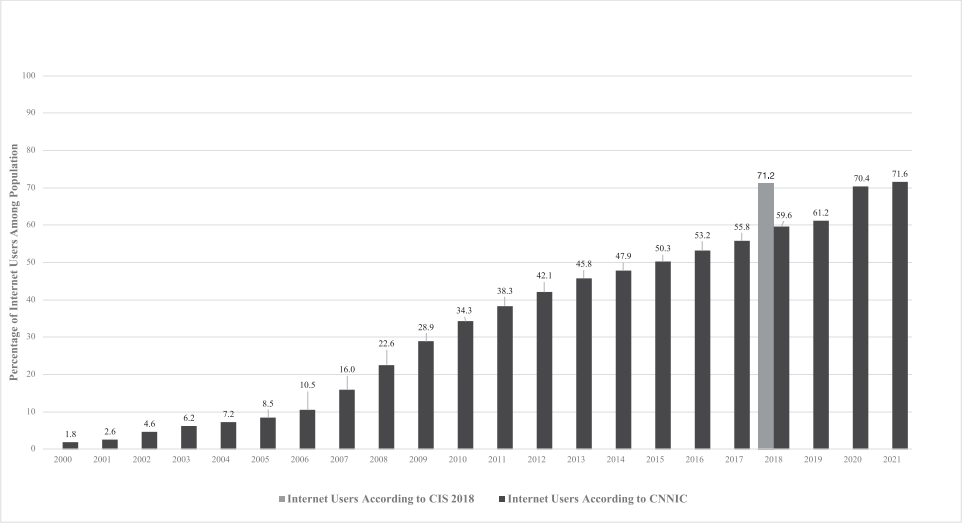


Figure 1. Percentage of Internet Users in China, 2000–2021
Source: CNNIC and CIS 2018.
Notes: Internet users expressed as percentages are taken from CNNIC reports. When percentages were not available (2000, 2001), we calculated the percentages of internet users using the real number of internet users provided by CNNIC reports and national population size from the National Bureau of Statistics. We used percentage of internet users by the end of each year. Mid-year (June) statistics were used when end of year data were not available (2019, 2021).

varies from province to province and the digital divide still exists."[38] "Due to a high correlation between internet development level and economic growth rate in different regions, the highest internet penetration rates are mainly seen in eastern provinces, while the lowest rate is mainly in south-western provinces."[39] These interpretations have become accepted wisdom in the study of the Chinese internet.

The CIS data also reveal a strong regional variation in internet use; however, there is no evidence in the CIS data of a relationship between economic development and internet penetration rates. Figure 2 displays internet use by province in 2018, with the darker grey representing higher percentages of internet use. Provinces in white were not sampled. In 22 out of 23 provinces, at least 60 per cent of the regional population uses the internet. Gansu, a landlocked province in central China, is an exception, with only 44 per cent. Anhui, landlocked and located in eastern China, has the highest percentage of internet users at 84 per cent. These patterns, however, reveal no correlation between the share of internet users and provincial economic development, nor that the eastern coastline has a larger share of internet users than other parts of China.

These patterns are further illustrated in Figure 3, where each province is plotted against its percentage of internet users based on CIS 2018 data on the y-axis and their 2017 gross regional product (GRP) in the x-axis. On the left-hand side of Figure 3, we see that many provinces with lower GRP – for example, Ningxia and Qinghai – have similar levels of internet penetration to Guangdong and Jiangsu. Western China has already caught up in internet penetration, despite its lower economic development.

One possible explanation for this observation is that internet penetration rates are not only shaped by "push" factors, such as infrastructure, but also by "pull" factors that affect ICT adoption rates at an individual level. One recent study on spatial variation in digital development across

38 CNNIC 2016, 48.
39 CNNIC 2017, 42. CNNIC reports contain internet penetration rates per province between 2013 and 2017. Unfortunately, no such data are available to compare with the period during which the CIS was conducted.

https://doi.org/10.1017/S0305741025000177 Published online by Cambridge University Press

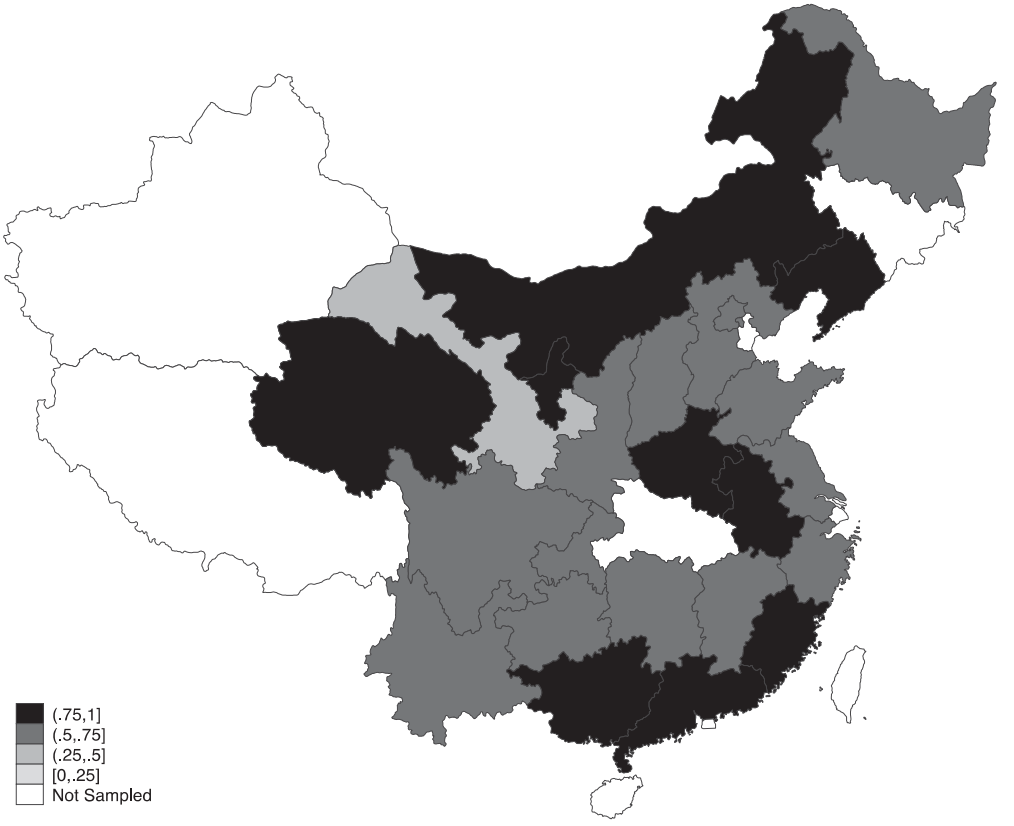


Figure 2. Percentage of Internet Users by Province, 2018
Source: CIS 2018.

China finds that not only does gross income have a significant impact but also the working age population ratio and secondary education enrolment ratio.[40] This suggests that, apart from infrastructure and financial resources, digital development also rests on the local population's attraction to and skills with digital technologies. Considering that CNNIC also samples school-aged individuals aged between 6 and 18, the digital divide portrayed by CNNIC statistics may reflect regional disparities in education resources, which affect the local population's skills with digital technologies, rather than infrastructure and economic growth.

In addition to variation across provinces, research on China's digital divide has also examined the urban–rural divide. CNNIC reports between 2013 and 2023 contain sections on the "Scale of internet users in rural areas." Since CNNIC constitutes official data, we assume that measurement in these sections relies on the official administrative classification of the location in which respondents were interviewed. However, Chinese official administrative classifications of rural and urban are not necessarily equivalent to urbanization in terms of population density. While many urban towns gained their urban status because of increasing population density, there are many towns with low population density that received their urban status because of a strategic choice made by local governments aimed at fostering new infrastructure projects and real estate development.[41] As a result, towns can vary significantly in population density. Our classification of urban and rural areas takes into account both the Chinese official administrative classifications and population density using

40 Song, Zhouying, Wang and Bermann 2020.
41 Chen et al. 2014; Chen, Davis and Landry 2017.

https://doi.org/10.1017/S0305741025000177 Published online by Cambridge University Press

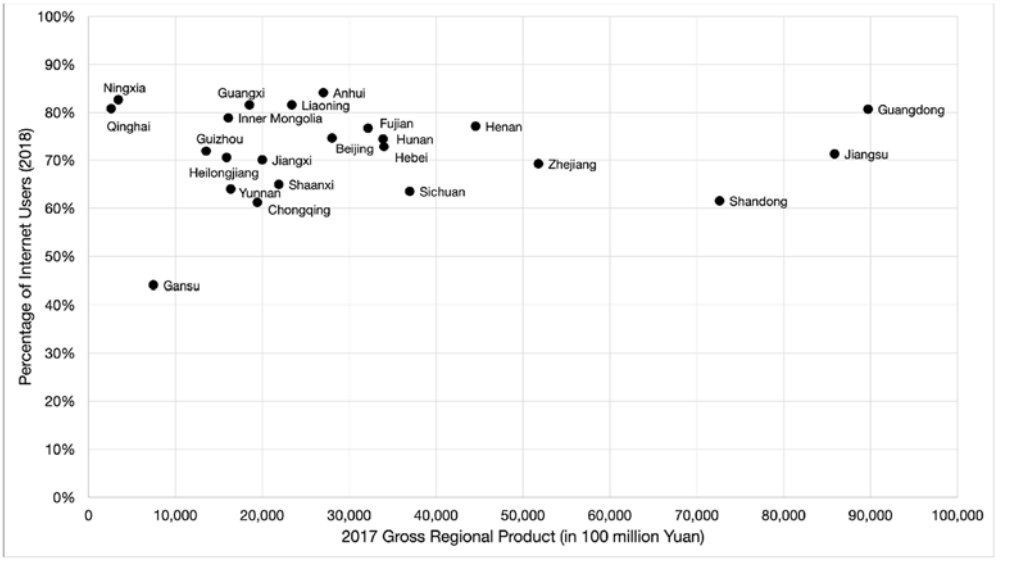


Figure 3. Percentage of Internet Users in 2018 and Growth Rate of 2017 Gross Regional Product, by Province
Source: National Bureau of Statistics 2018; CIS 2018.

data from WorldPop. We follow official classifications in classifying "rural town" (*xiang* 乡) as rural and "district" (*qu* 区) or "street" (*jiedao* 街道) as urban. We introduced population density to differentiate urban "towns" (*zhen* 镇) from rural towns – i.e. those that likely gained urban status because of the strategic choice of local governments. Towns from the upper quintile of cell density were coded as urban, while towns from the remainder were coded as rural.

These differing measurements lead to significantly differing understandings of the urban–rural digital divide. CNNIC concludes that "the internet penetration in rural areas rose to 34.0 per cent, but was still 35.4 per centage points lower than that in urban areas."[42] Thus, according to CNNIC, there remains a large gap between rural and urban areas, and this gap remains little changed over time, despite rising internet usage.[43]

By contrast, the CIS data reveal that rural China had almost caught up with cities by 2018, at least in terms of internet penetration. According to the CIS, in 2018, 66 per cent of people living in rural areas had access to the internet while 73 per cent of people living in urban areas were online. Although the gap in internet penetration is still significant, it is much narrower than the CNNIC statistics indicate. While the classification of *zhen* with lower population density as rural partly explains the narrower rural–urban digital divide, policy measures aimed at promoting digital development aided considerably in closing the rural and urban gap.[44] By the time the CIS was conducted, a large e-commerce firm no longer had to invest in increasing internet access to bring e-commerce to the countryside; government programmes had already achieved this. Instead, the firm focused on expanding e-commerce based on training and supporting digital skills.[45]

Overall, our comparison of official CNNIC data and the CIS data reveals different interpretations of the narrowing of the digital divide in China. While the 2018 CNNIC data confirm the common assumption that the majority of internet users tend to be located in more urban and economically

42 CNNIC 2018, 14.
43 CNNIC 2013; 2023.
44 In 2018, 63% of people living in *xiang* (rural areas according to the Chinese official classifications) were online, a significantly higher number than the reported 34% in CNNIC statistics.
45 Couture et al. 2021.

https://doi.org/10.1017/S0305741025000177 Published online by Cambridge University Press

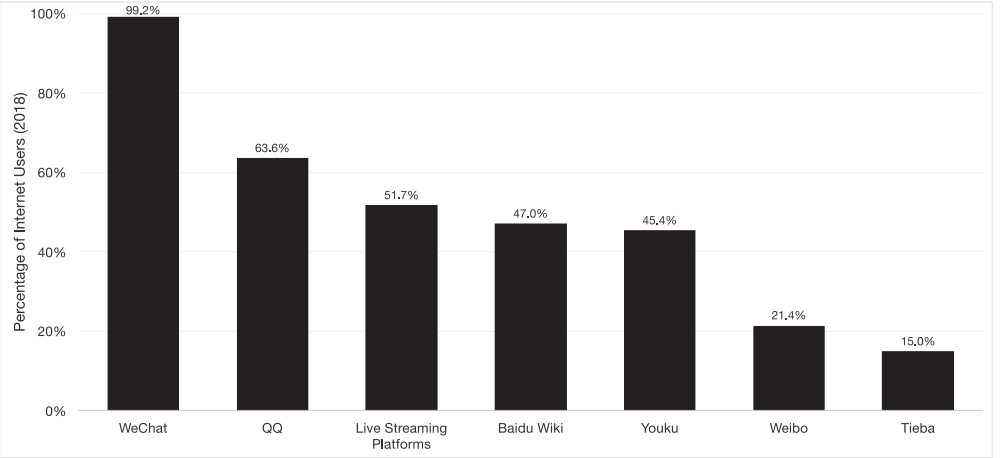


Figure 4. Percentage of Internet Users Using Social Media Platforms, 2018
Source: CIS 2018.

developed regions, the CIS data indicate that these urban–rural disparities are no longer as big as they were in the past. Instead, internet usage had already spread to the western inland and rural areas by 2018. This observation seems to be in line with studies that predicted that the digital divide in terms of internet access would likely decrease over time.[46]

Example 3: Understanding Social Media Use

Platform-specific information on social media use is important for scholarship on media consumption,[47] collective action,[48] public opinion,[49] crisis communication,[50] mental and physical health,[51] and identity,[52] among other topics. Many of these studies are based on the assumption that Weibo and WeChat are "the two main social networking sites in the country," according to business data from the platforms[53] or web traffic by Alexa or another similar web matrix.[54] Depending on which measurement was applied, scholars have come to different conclusions about the use of specific social media platforms.

CNNIC reports currently do not track social media use based on users by platform. Instead, they categorize social media based on functions, such as instant messaging, search engines or online news. For example, WeChat use is split between instant messaging and social networking. In a second step, instant messaging is reported not by platform but as a general category of internet use. While it is certainly important to understand why social media are used, too much uncertainty exists regarding the size of platform-specific user data in China to allow for systematic comparison across self-reported use of social media platforms. Thus, CIS respondents were given a list of ten platforms and asked whether they had accessed any of these platforms in the past three months. Results are displayed in Figure 4.

46 See, e.g., van Dijk and Hacker 2003.
47 Zhang and Guo 2021.
48 Zhao, Guan and Liang 2022.
49 Guo, Jing, and Hu 2023.
50 Zhao, Guan and Liang 2022.
51 Niu, Willoughby and Zhou 2021.
52 Jiang and Yang 2022.
53 Guo, Lei, and Chen 2021, 983.
54 Cosenza 2019. After 2021, these data indicate that WeChat is the most popular platform in China.

https://doi.org/10.1017/S0305741025000177 Published online by Cambridge University Press

According to CIS estimates, 99 per cent of internet users used WeChat in 2018, while QQ ranked as the second most widely used platform, with 64 per cent of internet users. In contrast, only 21 per cent of internet users were on the Twitter-like platform, Weibo. These figures reveal that Tencent dominates the social networks on the Chinese internet. In comparison, CNNIC reported lower percentages for social networking on WeChat Moments (83 per cent) and Qzone (59 per cent), but a significantly higher usage of Weibo (42 per cent).[55] However, these percentages do not provide conclusive information about the use of specific platforms by internet users.

Conclusion

This comparison of official CNNIC data and nationally representative CIS survey data illustrates how methodological differences in sampling design and measurement lead to substantially different conclusions about internet use in mainland China. These findings do, of course, have limitations. The most obvious is that the CIS was conducted at one moment in time, while CNNIC continues to survey internet users at regular intervals. Ideally, nationally representative surveys on internet use should be conducted more frequently to allow for longitudinal comparisons. However, given the recent challenges in conducting surveys in China, no such surveys have been conducted to our knowledge since 2018. In light of the reduced opportunities for researchers to collect nationally representative face-to-face surveys based on collaborations inside and outside China, we recommend that researchers remain aware of how sampling can lead to bias in conclusions drawn from reports. Most importantly, we found that official statistics significantly underestimate overall internet use rates. Our current estimate of bias is roughly 15 per cent. At the same time, CNNIC seems to significantly overestimate the regional and urban–rural digital divides. By December 2018, CNNIC estimated a 37 per cent divide between urban and rural internet penetration rates; our estimates suggest a 7 per cent divide. This difference is, to a large extent, the result of different classifications of rural and urban areas – one based on official definitions, the other based on scholarly work in the social sciences. The CIS data thus reveal that the Chinese internet has become more inclusive – at least in terms of access and especially for people living in rural and less developed regions. Similarly, different conceptualizations and measurements can also affect conclusions about platform usage. For instance, CNNIC underestimated WeChat and QQ use and overestimated Weibo use. We therefore caution against the use of CNNIC reports for research into the digital divide and platform usage rates.

Our findings have several implications regarding sourcing information about the Chinese internet. First, official statistics will most likely continue to be the most important data source, particularly for information about Tibet and Xinjiang where non-official data cannot be collected due to sensitivity issues. CNNIC already has become more transparent in its reporting, especially in its Chinese version. This positive development could be further supported with the provision of information on response rates, sample size, the size of overlap in mobile and landline samples, and other more detailed information about random sampling, weights and analysis. An online tool to calculate descriptive statistics along the lines of the one provided by the World Values Survey would likely improve our understanding of the data provided by CNNIC.

Our findings also have implications for alternative methods of data collection. Online surveys are becoming increasingly common in social science research on China (see Appendix Figure A1), partly owing to the challenges of conducting face-to-face surveys, as highlighted above. However, we remain cautious about using online surveys to draw conclusions about internet users. According to the CIS 2018, only 3 per cent of internet users reported participating in online surveys or other forms

55 CNNIC 2019. CIS also asked WeChat users whether they used Moments or Chats more often. To disseminate information, 45% reported using Chats, 19% used Moments more often, and 36% used both "about the same." For social networking, 71% preferred Chats, 3% preferred Moments and 25% used both equally. Overall, Chats is more frequently used for communication compared to Moments.

https://doi.org/10.1017/S0305741025000177 Published online by Cambridge University Press

of paid data collection and analysis, such as the Chinese versions of Mechanical Turk. These users tend to be younger, more educated, wealthier, and more likely to be migrant workers and located in eastern China. While the overall internet population has become more similar to the average Chinese, participants in online surveys have not. Although official statistics may not be entirely accurate, it is unfortunate that online surveys, too, likely offer a biased understanding of internet users in China.

Overall, official data on internet users should be interpreted in light of its sampling and measurement. In our view, online surveys cannot substitute for face-to-face surveys that are carefully designed to draw nationally representative conclusions on internet use. In the past, nationally representative surveys have been conducted as a joint effort between researchers located inside and outside of mainland China. As we have shown above, restricting the possibilities for data collection not only reduces information but also leads to biased understandings of the development of internet use. Thus, limiting data access for researchers is unlikely to truly improve our understanding of the Chinese internet: we need to seek truth from facts.

Acknowledgements. We would like to thank our partner in China for valuable support in survey development and fieldwork. We are also grateful to our advisory board members, Pierre Landry, Kent Jennings, Melanie Manion, Shen Mingming, Tang Wenfang and Wang Yuhua, whose dedication and guidance made the survey work possible. For superb research assistance, we are grateful to Felix Garten, Paxia Ksatryo and Dion Stevers. Finally, we thank Aofei Lü for writing a Chinese abstract that accurately captures the essence of the paper.

Funding statement. The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013)/ERC Grant Agreement No. [338478]. The Hertie School of Governance and Leiden University are both beneficiaries of the grant.

Competing interests. None.

References

Alpert, Mark I. 1971. "Identification of determinant attributes: a comparison of methods." *Journal of Marketing Research* **8**(2), 184–191.

Birney, Mayling. 2014. "Decentralization and veiled corruption under China's 'rule of mandates'." *World Development* **53**, 55–67.

Chen, Juan, Shuo Chen, Pierre F. Landry and Deborah S. Davis. 2014. "How dynamics of urbanization affect physical and mental health in urban China." *The China Quarterly* **220**, 988–1011.

Chen, Juan, Deborah S. Davis and Pierre F. Landry. 2017. "Beyond hukou reform: enhancing human-centered urbanization in China." Paulson Policy Memorandum, 17 February, https://paulsoninstitute.org.cn/wp-content/uploads/2017/02/PPM_Beyond-Hukou_Chen_Davis_Landry_English.pdf.

CNNIC (China Internet Network Information Center). 2013. "The 31st statistical report on internet development in China." China Internet Network Information Center, https://www.cnnic.com.cn/IDR/ReportDownloads/201302/P020130312536825920279.pdf.

CNNIC. 2016. "The 37th statistical report on internet development in China." China Internet Network Information Center, https://www.cnnic.com.cn/IDR/ReportDownloads/201604/P020160419390562421055.pdf.

CNNIC. 2017. "The 39th statistical report on internet development in China." China Internet Network Information Center, https://www.cnnic.com.cn/IDR/ReportDownloads/201706/P020170608523740585924.pdf.

CNNIC. 2018. "The 40th statistical report on internet development in China." China Internet Network Information Center. https://www.cnnic.com.cn/IDR/ReportDownloads/201807/P020180711387563090220.pdf.

CNNIC. 2019. "The 43rd statistical report on internet development in China." China Internet Network Information Center, https://www.cnnic.com.cn/IDR/ReportDownloads/201911/P020191112538996067898.pdf.

CNNIC. 2023. "The 52nd statistical report on internet development in China." China Internet Network Information Center, https://www.cnnic.com.cn/IDR/ReportDownloads/202311/P020240325599899189635.pdf.

Cosenza, Vincenzo. 2019. "World map of social networks 2019," https://vincos.it/world-map-of-social-networks/. Accessed 10 November 2020.

Couture, Victor, Benjamin Faber, Yizhen Gu and Lizhi Liu. 2021. "Connecting the countryside via e-commerce: evidence from China." *American Economic Review: Insights* **3**(1), 35–50.

Dutton, William H., Everett M. Rogers and Suk-Ho Jun. 1987. "Diffusion and social impacts of personal computers." *Communication Research* **14**(2), 219–250.

https://doi.org/10.1017/S0305741025000177 Published online by Cambridge University Press

Fieldhouse, Ed, Jane Green, Geoffrey Evans, Hermann Schmitt, Cees van der Eijk, Jon Mellon and Chris Prosser. 2019. "British election study, 2017: face-to-face post-election survey." UK Data Service, SN: 8418, DOI: 10.5255/UKDA-SN-8418-1, https://www.britishelectionstudy.com/data-object/2017-face-to-face/.

Fisher, Robert J. 1993. "Social desirability bias and the validity of indirect questioning." *Journal of Consumer Research* **20**(2), 303–315.

Greitens, Sheena Chestnut, and Rory Truex. 2020. "Repressive experiences among China scholars: new evidence from survey data." *The China Quarterly* **242**, 349–375.

Guo, Jing, and Yang Hu. 2023. "Does social media use polarize or depolarize political opinion in China? Explaining opinion polarization within an extended communication mediation model." *Social Media + Society* **9**(3), https://doi.org/10.1177/20563051231196899.

Guo, Lei, and Hsuan-Ting Chen. 2021. "The impact of social media on civic engagement in China: the moderating role of citizenship norms in the citizen communication mediation model." *Journalism and Mass Communication Quarterly* **99**(4), 980–1004.

Huang, Yasheng. 2008. "A great reversal." In *Capitalism with Chinese Characteristics: Entrepreneurship and the State*. Cambridge: Cambridge University Press, 175–232.

Jiang, Yao, and Fan Yang. 2022. "Bridge of rainbow: association between internet-based social media use and homosexuality inclusion in China." *Frontiers in Psychology* **13**, 1–11.

King, Gary, Christopher Murray, Joshua A. Salomon and Tandon Ajay. 2004. "Enhancing the validity and cross-cultural comparability of measurement in survey research." *American Political Science Review* **98**(1), 191–207.

Landry, Pierre F., and Mingming Shen. 2005. "Reaching migrants in survey research: the use of the global positioning system to reduce coverage bias in China." *Political Analysis* **13**(1), 1–22.

Lei, Xuchuan, and Jie Lu. 2017. "Revisiting political wariness in China's public opinion surveys: experimental evidence on responses to politically sensitive questions." *Journal of Contemporary China* **26**(104), 213–232.

Livingstone, Sonia, and Ellen Helsper. 2007. "Gradations in digital inclusion: children, young people and the digital divide." *New Media and Society* **9**(4), 671–696.

Manion, Melanie. 2010. "A survey of survey research on Chinese politics." In Allen Carlson, Mary E. Gallagher, Kenneth Lieberthal and Melanie Manion (eds.), *Contemporary Chinese Politics: New Sources, Methods, and Field Strategies*. New York: Cambridge University Press, 181–199.

Marshall, Eliot. 1993. "US may renew collaboration after China relents on data." *Science* **261**(5122), 677.

National Bureau of Statistics. 2015. *China Statistics Yearbook 2015*. Beijing: China Statistics Press.

National Bureau of Statistics. 2018. *China Statistics Yearbook 2018*. Beijing: China Statistics Press.

National Bureau of Statistics. 2021. "Diqici quanguo renkou pucha zhuyao shuju qingkuang" (Main data of the seventh national census), http://www.stats.gov.cn/tjsj/zxfb/202105/t20210510_1817176.html. Accessed 23 February 2022.

Naughton, Barry. 2018. "The urban–rural divide and Chinese-style urbanization." In *The Chinese Economy: Adaptation and Growth* (2nd ed.). Cambridge, MA: MIT Press, 127–152

Niu, Zhaomeng, Jessica Willoughby and Rongting Zhou. 2021. "Associations of health literacy, social media use, and self-efficacy with health information-seeking intentions among social media users in China: cross-sectional survey." *Journal of Medical Internet Research* **23**(2), e19134.

Norris, Pippa. 2001. *Digital Divide: Civic Engagement, Information Poverty, and the Internet Worldwide*. Cambridge: Cambridge University Press.

Oi, Jean C. 1999. "Two decades of rural reform in China: an overview and assessment." *The China Quarterly* **159**, 616–628.

Piotrowski, Suzanne J., Yahong Zhang, Weiwei Lin and Wenxuan Yu. 2009. "Key issues for implementation of Chinese open government information regulations." *Public Administration Review* **69**(s1), S129–S135.

Plekhanov, Dmitriy. 2017. "Quality of China's official statistics: a brief review of academic perspectives." *The Copenhagen Journal of Asian Studies* **35**(1), 76–101.

Roberts, Margaret E. 2018. *Censored: Distraction and Diversion Inside China's Great Firewall*. Princeton, NJ: Princeton University Press.

Robinson, Darrel, and Marcus Tannenberg. 2019. "Self-censorship of regime support in authoritarian states: evidence from list experiments in China." *Research and Politics* **6**(3), 1–9.

Solinger, Dorothy J. 1999. *Contesting Citizenship in Urban China: Peasant Migrants, the State, and the Logic of the Market*. Berkeley, CA: University of California Press.

Song, Wei. 2008. "Development of the internet and digital divide in China: a spatial analysis." *Intercultural Communication Studies* **17**(3), 20–43.

Song, Zhouying, Chen Wang and Luke Bergmann. 2020. "China's prefectural digital divide: spatial analysis and multivariate determinants of ICT diffusion." *International Journal of Information Management* **52**, 102072.

Stockmann, Daniela. 2010. "Information overload? Collecting, managing, and analyzing Chinese media content." In Allen Carlson, Mary E. Gallagher, Kenneth Lieberthal and Melanie Manion (eds.), *Contemporary Chinese Politics: New Sources, Methods, and Field Strategies*. New York: Cambridge University Press, 107–126

Stockmann, Daniela, Ashley Esarey and Jie Zhang. 2018. "Who is afraid of the Chinese state? Evidence calling into question political fear as an explanation for overreporting of political trust." *Political Psychology* **39**(5), 1105–21.

https://doi.org/10.1017/S0305741025000177 Published online by Cambridge University Press

Stockmann, Daniela, and Ting Luo. 2018. "China internet survey 2018," https://chinainternetsurvey.net/.
Stockmann, Daniela, and Ting Luo. Forthcoming. *Governing Digital China*. New York: Cambridge University Press.
Stromseth, Jonathan R., Edmund J. Malesky and Dimitar D. Gueorguiev. 2017. *China's Governance Puzzle: Enabling Transparency and Participation in a Single-party State*. Cambridge: Cambridge University Press.
Su, Chris Chao, Jun Liu and Baohua Zhou. 2020. "Two levels of digitalization and internet use across Europe, China and the US." *International Journal of Communication* **14**, 5838–59.
Tang, Wenfang. 2016. *Populist Authoritarianism: Chinese Political Culture and Regime Sustainability*. Oxford: Oxford University Press.
Tourangeau, Roger, and Ting Yan. 2007. "Sensitive questions in surveys." *Psychological Bulletin* **133**(5), 859–883.
Truex, Rory. 2017. "Consultative authoritarianism and its limits." *Comparative Political Studies* **50**(3), 329–361.
Tsai, Lily L. 2010. "Quantitative research and issues of political sensitivity in rural China." In Allen Carlson, Mary E. Gallagher, Kenneth Lieberthal and Melanie Manion (eds.), *Contemporary Chinese Politics: New Sources, Methods, and Field Strategies*. New York: Cambridge University Press, 246–265.
Van Deursen, Alexander J.A.M., and Jan A.G.M. van Dijk. 2014. "The digital divide shifts to differences in usage." *New Media and Society* **16**(3), 507–526.
Van Dijk, Jan, and Kenneth Hacker. 2003. "The digital divide as a complex and dynamic phenomenon." *The Information Society* **19**(4), 315–326.
Wallace, Jeremy L. 2016. "Juking the stats? Authoritarian information problems in China." *British Journal of Political Science* **46**(1), 11–29.
Wallace, Jeremy L. 2022. *Seeking Truth and Hiding Facts: Information, Ideology, and Authoritarianism in China*. Oxford: Oxford University Press.
WorldPop. 2022. "Open spatial demographic data and research," https://www.worldpop.org/. Accessed 23 February 2022.
Zhang, Yiyan, and Lei Guo. 2021. "'A battlefield for public opinion struggle': how does news consumption from different sources on social media influence government satisfaction in China?" *Information, Communication and Society* **24**(4), 594–610.
Zhao, Xin, Mengfei Guan and Xinya Liang. 2022. "The impact of social media use on online collective action during China's COVID-19 pandemic mitigation: a social identity model of collective action (SIMCA) perspective." *International Journal of Communication* **16**, 85–106.
Zhou, Baohua, Chris Chao Su and Jun Liu. 2021. "Multimodal connectedness and communication patterns: a comparative study across Europe, the United States, and China." *New Media and Society* **23**(7), 1773–97.

Appendix

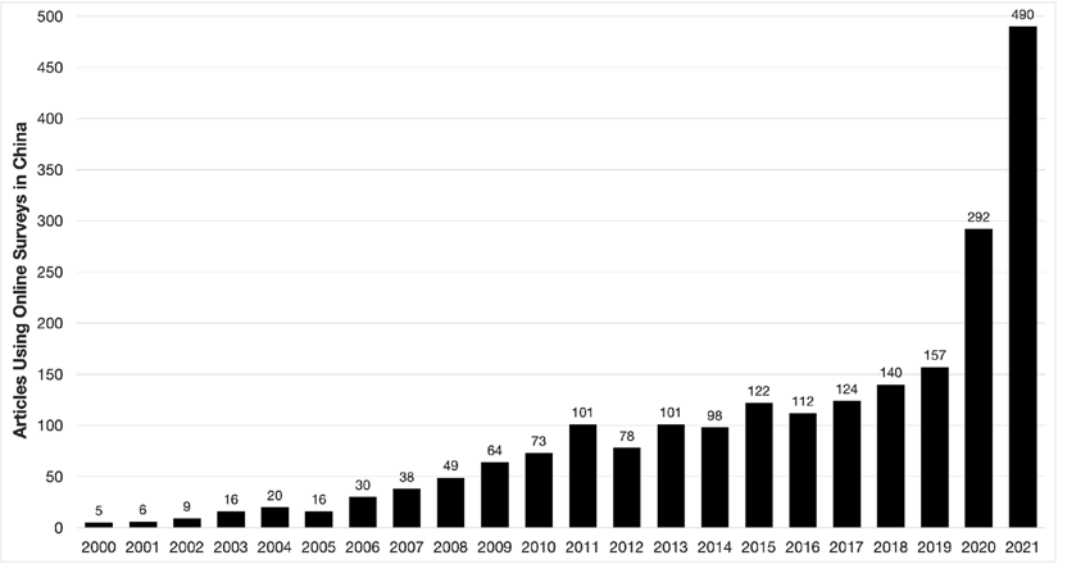


Figure A1. Frequencies of Articles Using Online Surveys in China, 2000–2021
Source: Social Science Citation Index.
Note: 12 December 2024 version.

https://doi.org/10.1017/S0305741025000177 Published online by Cambridge University Press

Daniela STOCKMANN is director of the Center for Digital Governance, and professor of digital governance, at the Hertie School. Her 2013 book, *Media Commercialization and Authoritarian Rule in China*, received the 2015 Goldsmith Book Prize from the Harvard Kennedy School. Her current research focuses on the interaction between government, platform firms and citizens in the area of social media governance. She studies these interactions both in China and in Europe. Her forthcoming book, *Governing Digital China* (with Ting Luo, Cambridge University Press), challenges top-down notions of digital governance and explores the logic of popular corporatism, highlighting bottom-up influences of China's largest platform firms.

Ting LUO is an associate professor in government and artificial intelligence at the University of Birmingham. Her research focuses on the intersection of technology and governance, with particular emphasis on digital governance.

Cite this article: Daniela Stockmann and Ting Luo. 2025. "Surveying Internet Users in China: Comparing Representative Survey Data with Official Statistics." *The China Quarterly* **262**, 531–547. https://doi.org/10.1017/S0305741025000177

https://doi.org/10.1017/S0305741025000177 Published online by Cambridge University Press