# Consultation and Selective Censorship in China

**Dimitar D. Gueorguiev**, Syracuse University
**Edmund J. Malesky**, Duke University

Conventional theories of authoritarianism view the need to suppress criticism as a key function of censorship. In a 2013 article, King, Pan, and Roberts challenged this wisdom by arguing that the paramount goal of censorship is defusing collective action, not silencing dissent. After accounting for collective action potential, they argue that criticism has no bearing on censorship. In this research note, we point out that a significant portion of sampled posts in King et al.'s analysis coincided with state-led consultation campaigns that were aimed at soliciting critical public input on policy proposals. This introduces the potential for bias by combining solicited and unsolicited criticism under the generic title of criticism. After reanalyzing King et al.'s aggregate data, studying Chinese censorship directives, and offering a statistically guided thought experiment, we conclude that a more conservative version of their original thesis is in order.

What do authoritarian regimes censor? In China, censors routinely suppress the voice of dissidents, lawyers, intellectuals, and prominent personalities (Dimitrov 2017), while information authorities monitor and coerce publishers to self-censor (*Economist* 2017). All the while, algorithms and moderators scrub the Chinese internet of problematic keywords and even the unseemly image of "Winnie the Pooh," presumably because he shares an amusing resemblance to current General Party Secretary Xi Jinping (McDonell 2017).

Such targets are consistent with conventional theories on authoritarianism, which view the need to suppress dissent as the principal objective of censorship (Geddes and Zaller 1989). In an important article, King, Pan, and Roberts (2013) challenged this wisdom, positing that the goal of authoritarian censorship is not stifling criticism but defusing collective action potential (CAP), such as calls for organized protests or other forms of group expression. The data-intensive article has been justifiably influential, collecting over 1,050 citations in five years and appearing on syllabi around the world.

While sympathetic to King et al. in spirit, we suggest that dismissing criticism as a determinant of censorship is unwarranted. Specifically, we point out that 16% of their 87 content topics (and 78% of policy-related topics that also happen to be the least censored) coincided with or directly followed from state-led consultation campaigns—a procedure whereby policy makers publicize select policy proposals for debate and public comment (Chen and Xu 2017; Truex 2014). By not taking this nuance into account, researchers risk misinterpreting and underestimating the limits to critical discourse in China's single-party state. Our best attempts at replicating and expanding on King et al.'s original findings suggest that a more conservative version of their thesis is in order.

## REVISITING SELECTIVE CENSORSHIP

There are both weak and strong readings of King et al.'s main thesis. The weak reading is that, all else equal, CAP increases the likelihood of censorship. The strong reading is that the regime is far more likely to censor content involving CAP but is indifferent to criticism. According to King et al., evidence for the strong interpretation is unambiguous: "posts are censored if they are in a topic area with CAP and not otherwise. Whether or not the posts are in favor of the government, its leaders, and its policies has no measurable effect on the probability of censorship" (2013, 338). Evidence for the strong claim is presented in their figure 9, reproduced here as figure 1, which shows no difference in censorship rates for critical posts compared to those in support of the regime.

Dimitar D. Gueorguiev (ddgueorg@syr.edu) is an assistant professor of political science at Syracuse University, Syracuse, NY 13244. Edmund J. Malesky (ejm5@duke.edu) is a professor of political science at Duke University, Durham, NC 27708.

**1539**

Figure 1. Reproduction of figure 9 from King et al. (2013, 337)

We argue that the strong reading does not consider the varieties of criticism that exist within a controlled social medium. In particular, it does not account for the fact that the Chinese state often solicits public criticism when formulating policy proposals, much like the "prompted criticism" featured in Soviet-era, state-run newspapers (Fitzpatrick 1996; Inkeles and Geiger 1953). This kind of criticism is fundamentally different from unsolicited critiques, such as those targeting political figures or regime ideology. Moreover, we argue that King et al.'s sampling strategy, which focused on national-level bursts of online discussion, is particularly likely to have scooped up solicited criticism emanating from state-led consultation campaigns.

We believe that these features of King et al.'s research design are driving the null effect in the right panel of figure 1 (non-CAP events). If consultation-related posts were removed from the analysis, we should observe significantly more censorship of critical than supportive posts in that category.

## SOLICITED VERSUS UNSOLICITED CRITICISM

Each year, China subjects hundreds of draft laws and regulations to public review, at both the national and provincial levels.[1] Scholars describing this type of engagement broadly refer to it in terms of "consultative authoritarianism" (He and Warren 2011; Truex 2014). During the consultation process, the regime publicizes policy proposals and citizens weigh in, often in a critical fashion (Horsley 2009). The best argument for why a nondemocratic regime might encourage such behavior is consistent with King et al.'s logic for why criticism is uncensored: it is "an effective governmental tool in learning how to satisfy, and ultimately mollify, the masses" (King et al.

---

1. The most common form of review is Notice and Comment. In 2010–11, over 100 national and 300 provincial policy proposals were subjected to this process (Stromseth, Malesky, and Gueorguiev 2017, 175, 210–21).

2013, 339). Moreover, there is evidence that consultation works. Gueorguiev (2014), for instance, finds that laws and regulations that undergo consultation are less likely to be repealed or amended, indicating that learning contributes to policy stability.

For consultation to facilitate learning, however, there must be reasonably open debate; that is, policy makers must tolerate critical feedback. For instance, consultations over the 2006 draft Labor Contract Law led to a heated and public exchange between the state-run labor union, more business-friendly segments of the government, and China's upper house of congress (Stromseth et al. 2017, 177–79). More recently, a seemingly minor proposal to lower the legal age for marriage unleashed a furor of public criticism, much of it aimed at the regime's broader intrusions into family planning (Koetse 2017). We refer to this type of criticism as "solicited" and argue that it is less likely to be censored precisely because it is sanctioned by and serves the interests of the regime.

By contrast, criticism of the regime and its policies that does not arise through public consultation is "unsolicited." Unsolicited criticism can be general or specific, benign or sensitive, can discuss individual policies or the regime as a whole and, on occasion, can even impugn the reputation of individual leaders (Esarey and Xiao 2008). The critical distinction is that unsolicited criticism emerges outside the regime's preferred information-gathering channels and is therefore more likely to be censored. A blunt measure of criticism that does not account for consultation is therefore incapable of capturing the countervailing effects of solicited and unsolicited criticism, which biases toward a null effect of criticism on censorship.

## INCORPORATING CONSULTATION INTO KING AND COLLEAGUES' RESEARCH DESIGN

How likely is it that consultation-inspired, solicited criticism is influencing King et al.'s conclusions? Part of the answer lies in the data-generating process underpinning their analysis. The authors first use expert opinion and commentary to identify 85 topic areas, ranging across different levels of sensitivity. Over six months in 2011, they collected all posts in the time window using keywords for these topics, yielding roughly 3.7 million nonspam posts. They repeated these steps for time periods outside of the six-month window, raising the total number collected to 11.4 million. King et al. then repeatedly revisited each post to determine if it was censored.

Stratifying across topic sensitivity, King et al. randomly sampled 127,283 posts for further analysis. Within each topical sample, they then detected 87 bursts in volume, described as "sharp spikes in [posting] volume around specific subjects," relative to the rest of the six-month period (2013, 331). The authors later identified the external "event" associated with each

burst and coded each burst event into five mutually exclusive categories: (i) CAP, (ii) criticism of the censors, (iii) pornography, (iv) government policies, and (v) other news. Their headline inference about CAP is derived from the burst-level behaviors of censors.

The design is creative, but the sampling strategy has problematic implications for the analysis. As the authors note on page 330, and expand on in King, Pan, and Roberts (2014), "volume bursts" account only for topics that break through pervasive self-censorship and keyword blockers. Moreover, because their sampling frame considers the entirety of China, not only are such topics rare, but they must be broad enough in subject matter in order to capture an extremely diverse national audience. Put differently, this means that the "bursty" topics picked up by King et al. either must have taken censors by surprise, as in the case of the tragic Fukushima nuclear disaster in 2011, or were allowed to bypass them altogether, as in the state-initiated debate on amending the income tax threshold that same year.[2]

This selection process raises two concerns regarding King et al.'s findings, selection bias and confounding. First, the selection of topics and filtering on high-visibility bursts likely induced greater selection of posts related to government-led consultation campaigns (aimed at raising public awareness and input). These national-level conversations are more likely to have generated spikes in internet activity than complaints about local-level officials or everyday frustrations with government policy. Second, a lack of censorship within consultation-related bursts is likely to have been presanctioned by regime leaders, leading to conflation of solicited and unsolicited criticism under

the generic title of *Criticism*. We suggest that properly accounting for China's consultative approach to policy making would reveal critical differences in how China approaches solicited and unsolicited criticism. More precisely, we offer two hypotheses for non-CAP topics:

**H1.** Critical posts following from consultation (solicited) are less likely to be censored than unsolicited critical posts.

**H2.** Critical posts not following from consultation (unsolicited) are more likely to be censored than noncritical posts.

## EMPIRICAL ANALYSIS OF CRITICISM AND CENSORSHIP

Ideally, we should be able to take King et al.'s data, add a variable accounting for consultation, and estimate the censorship likelihood for unsolicited criticism. Unfortunately, their Criticism variable is not coded at the burst level. To generate their findings (reproduced in fig. 1), King et al. hand-code a smaller sample of posts as (i) critical (of the state), (ii) supportive, (iii) or irrelevant. They use this as training data for an extension of Hopkins and King (2010) to estimate the aggregate fraction of posts that are critical, supportive, or irrelevant in the censored subset of their sample, and then again for uncensored posts.[3]

Without a direct measure of criticism at the post or burst level, we probe our theory with three suggestive tests. First, we examine the relationship between consultation and censorship magnitude using replication data provided at the aggregate burst level. Second, we examine an external sample of leaked censorship directives to see if criticism solicited during consultation is less prone to censorship. Third, we perform a statistically guided thought experiment of what would happen if it were possible to rerun the King et al. analysis but drop critical posts occurring through solicited channels.

### Burst-level analysis

In our first test, we review each of the 18 bursts categorized as "Policy" in King et al.'s burst-level replication data set. Our review, summarized in table 1, finds that 14 of these bursts (78%) coincided with or directly followed from state-initiated calls for public input regarding ongoing policy debates. While

---

2. In a follow-up piece, King et al. (2014) account for unobserved heterogeneity in the 2013 observational study by using an experiment to randomly assign pro- and antigovernment positions in online posts. While reasonable for validating their findings on CAP, the King et al. experiment is insufficient to resolve our concerns about criticism, as they continue to employ the same topic sampling strategy: "we scoured the news and selected ongoing collective action events and non–collective action events about which there was a volume burst in social media discussion" (King et al. 2014, 1251722-4). Just as in their 2014 article, this means that they likely captured topics, like Xi Jinping's famous steamed bun eating stunt, on which the regime intended to generate public attention and discussion (see our app. B for discourse summaries on policy-related topics in King et al. [2014]). Position-taking sentiments within such topics are not synonymous with genuine criticism of the regime. In some cases, the "antigovernment" message describes displeasure with the policy without referencing government at all. For instance, the antigovernment posts for the rental tax policy treatment "criticize the tax saying it will increase already high rental taxes as landlords will push the tax onto renters" (King et al. 2014 supplementary materials, p. 3, available at http://www.sciencemag.org/content/345/6199/1251722/suppl/DC1). Moreover, due to the authors' burst selection strategy, criticisms that were suppressed during automatic review (1251722-2) or regionally concentrated cases of criticism never enter the burst sample and therefore are not considered in the experiment.

---

3. The authors provided us the event name, category, and a measure of average censorship magnitude calculated for the 87 bursts in their analysis. Due to the ReadMe software's probabilistic approach, however, these data did not contain measures of criticism at the post or burst level. See https://goo.gl/Mqu7yQ for the article's Dataverse page and web app. C for the burst-level spreadsheet.

Table 1. Policy Bursts Captured by King et al. (2013)

| | Observation Period | Consultation |
|---|---|---|
| 01. Reports about exchange rate, US Treasury Secretary pushes China on yuan | 1/13/11 | N/A |
| 02. US-China summit | 1/19/11 | N/A |
| 03. Wen Jiabao signs new regulations related to land and housing demolition | 1/19–1/22/11 | Yes |
| 04. Reports of news real estate tax pilot in Chongqing and Shanghai | 1/31–2/11/11 | Yes |
| 05. Minister of Health releases report on improving health access | 2/19/11 | Yes |
| 06. Clean-up campaign announced for lead acid batteries | 2/22/11 | Yes |
| 07. Reports that China will accelerate real tax pilots | 2/28/11 | Yes |
| 08. National People's Congress session focuses on health care reform | 3/5–3/16/11 | Yes |
| 09. Finance Ministry announces education reforms (migrant children) | 3/8/11 | Yes |
| 10. Adjustments to the One Child Policy discussed at the Two Meetings | 3/8–3/9/11 | Yes |
| 11. China puts nuclear program on hold after Japanese earthquake | 3/13–3/30/11 | N/A |
| 12. Reports that China delays smoking ban implementation | 3/23–3/25/11 | Yes |
| 13. NDRC raising price of oil | 4/7/11 | Yes |
| 14. Smoking ban comes into effect 5/1 | 4/30–5/12/11 | Yes |
| 15. Talk about letting electricity prices increase in June, during peak usage | 5/18–5/31/11 | Yes |
| 16. China places quota on rare earths | 5/20/11 | No |
| 17. China Environment Ministry issues new rules on lead/heavy metal pollution | 5/24–5/26/11 | Yes |
| 18. Reports of new indoor smoking ban | 5/31/11 | Yes |

we do not know what proportion of posts within consultation bursts directly address the policy debates, we do know that government reports frequently cite online debate and critical comments as evidence of a robust consultative process (see web app. A).

Figure 2, an extended version of King et al.'s figure 4, illustrates the share of censored posts, by burst. We depict bursts involving ongoing consultation with hollow circles and topics with CAP in squares. Consistent with our theory, consultation bursts are very unlikely to face censorship. The most censored burst—a pilot urban real estate tax—scores about 10 on censorship magnitude, but the modal consultation burst's score is less than zero, implying less censorship than the baseline in nonsensitive periods.[4] This fact only underscores our argument that policy discussion is actively encouraged, but only on consultation-related topics. By contrast, bursts coded as CAP face censorship magnitudes of up to 20 times greater than average magnitude across the entire sample.

In table 2, we go further by estimating the independent effect of consultation on censorship magnitude. Model 1 reports the burst-level regression of censorship magnitude on consultation, while model 2 controls for CAP. Both models employ

ordinary least squares estimates with robust standard errors to account for heteroskedasticity, as variance within CAP bursts is lower than in non-CAP bursts. We find that censorship in consultation bursts is about 5% below other bursts, a result that is significant at the .05 level.

## Validation using leaked directives

To validate that consultation topics were indeed spared from censorship, we reached out to China scholars in possession of leaked censorship directives, which overlap with King et al.'s six-month observation period. These logs, originally leaked to the Committee to Protect Journalists by an employee at China's second-largest social network site Sina Weibo (Miller 2017), document censorship instructions given to human censors employed by Sina, in its Tianjin office. The instructions contained in these logs come from a variety of sources and provide explicit instructions on how to manage online conversation. Blake Miller subsequently organized the logs into a database, allowing us to conduct keyword searches to see if any of the topics identified in table 1 were mentioned in the censorship directives.

On the basis of a review of over 10,000 directive logs, we can confirm that China did not issue a single directive for censoring critical commentary on consultation-related policies. Indeed, as panel A of table 3 reveals, our liberal keyword queries recovered only 10 censorship directives with any related content. After going through the text of each directive to make sure that the content was indeed related to the bursts identified

---

4. *Censorship magnitude* is defined as the share of posts censored during a volume burst minus the baseline share of censored posts that occur outside a volume burst (King et al. 2013, 332). This number can take on negative values if the censorship is less than the baseline (somewhere between 3% and 5%; 333n6).

Figure 2. Censorship magnitude for bursts with and without consultation. *Solid line*, mean censorship in sample; *dashed line*, mean censorship of policy and news without consultation.

by King et al. and reviewed in panel A, a process we refer to as "positive ID," we found none to be directly associated. It is worth noting that directives 3473, 3476, and 3522 were related to the National Development Reform Commission (NDRC) decision to raise fuel prices (see web app. A, sec. A.13), but censorship arrived only a week after the identified burst, when price hikes resulted in protests by taxi and lorry drivers. Thus, King et al.'s own coding strategy would likely have identified posts related to the drivers' protests as "CAP," not as "policy," because they involved protest. In other words, King et al.'s burst algorithm detected the nonsensitive discussion surrounding the policy announcement, but not the sensitive discussion that followed, presumably because censors were effective in suppressing it.

One drawback to our positive ID findings, summarized in panel A of table 3, is that they cover only consultation-related policy discussions. A more thorough examination of our theory, however, is that nonconsultation policy discussions are more likely to be censored. To see if this was the case, we conducted an additional, general query for policy discussions that did not involve consultation, with keywords for legislation, regulations, statutes, and guidelines. Looking within the month of May 2011, the final month of King et al.'s analysis, we identified eight possibly related directives. Consistent with hypothe-

sis 2, three of those discussions (37.5%)—concerning state secrets, bilingual education, and internet access—were flagged for censorship (see panel B of table 3). In short, on the basis of leaked censorship directives, policy-related discussions that

Table 2. Censorship and Consultation

|  | Consultation (1) | CAP Control (2) |
|---|---|---|
| Consultation | −.084** | −.048** |
|  | (.023) | (.021) |
| Collective action |  | .213** |
|  |  | (.063) |
| Constant | .052** | .016 |
|  | (.020) | (.018) |
| Observations | 87 | 87 |
| $R^2$ | .040 | .251 |
| Root mean squared error | .040 | .251 |
| BP test | 7.74 | 14.56 |
|  | (.005) | (.0001) |

Note. BP Test is the $\chi^2$ and (*p*-value) from the Breush-Pagan/Cook-Weisberg test for heteroskedasticity. Robust standard errors are in parentheses.
* $p < .10$.
** $p < .05$.

Table 3. Censorship Directives: Policy Queries

| Directive | Query | Date | Positive ID |
|---|---|---|---|
| | | | A. Consultation Policy Queries |
| 3390 | Demolition | 2/18/11 | 0: censors instructed not to provoke demolition activist, Zhu Fuxiang |
| 3468 | Demolition | 4/13/11 | 0: censor discussion of antidemolition protests (Fuyou Street) |
| 3472 | Demolition | 4/15/11 | 0: see above |
| 3480 | Demolition | 4/15/11 | 0: see above (log 3468) |
| 3482 | Demolition | 4/15/11 | 0: see above (log 3468) |
| 3473 | Fuel prices | 4/15/11 | 1: after King et al.'s 4/7/2011 observation; taxi protests |
| 3476 | Fuel prices | 4/15/11 | 1: after King et al.'s 4/7/2011 observation; taxi protests |
| 3489 | Demolition | 4/18/11 | 0: see above (log 3468) |
| 3522 | Fuel prices | 4/20/11 | 1: after King et al.'s 4/7/2011 observation; lorry driver protest in Shanghai |
| 3677 | Demolition | 5/16/11 | 0: Shanghai (Wujingzhen) protest against demolition/relocation plan |
| | | | B. Nonconsultation Policy Queries |
| 3648 | Policy | 5/10/11 | 0: sanitizing discussion of left/right factional infighting in the party |
| 3717 | Policy | 5/17/11 | 1: limit sensitive discussion about the new state secrets policy |
| 3806 | Policy | 5/28/11 | 1: limit discussion on bilingual education policy, academic's petition |
| 3822 | Policy | 5/27/11 | 1: (see above) come up with keywords to block academic's petition |
| 3653 | Guideline | 5/10/11 | 0: censor discussion of Jiangxi student's rape by foreign students |
| 3750 | Guideline | 5/24/11 | 0: general encouragement to step up antirumor efforts |
| 3768 | Guideline | 5/24/11 | 0: general encouragement to step up antirumor efforts |
| 3649 | Plan | 5/10/11 | 1: censor discussion on registration policy for international internet servers |

did not involve consultation are more likely to be censored than those arising within the regime's preferred information-gathering approaches.

### Thought experiment

While the leaked directives analysis demonstrates that consultation topics incur less censorship, it remains unclear what this means for King et al.'s findings on criticism overall. Unfortunately, this is difficult to test directly because their content analysis approach generates a corpus-level probability, but not a post-level or burst-level measurement of criticism.[5] Thus, we perform a statistically guided thought experiment in web appendix D, where we simulate what might happen to results in the right-hand (non-CAP) panel of figure 1 if consultation posts were excluded from the analysis.

Using a difference-in-proportions approach, and assuming an $N$ of 2,000 posts for figure 1, we find that dropping any more than 97 consultation-related posts produces a significant difference in censorship propensity between the remaining critical

and noncritical posts.[6] This is consistent with hypothesis 2, that China is likely to censor critical posts when they are unsolicited.

### DISCUSSION

This note revisits the issue of selective censorship under autocracy explored in King et al.'s influential 2013 paper. Our review shows that a substantial portion of their sample includes topics that coincided with or directly followed from state-initiated calls for public input regarding open policy debates. Our analysis of King et al.'s replication data and censorship directives from the Chinese government demonstrates that retaining these topics as evidence of uncensored criticism is potentially problematic.

To be clear, our review does not challenge the broader conclusions of King et al.'s analysis, that is, that nondemocratic regimes fear collective action and are thus likely to suppress it in various ways, including censorship. Moreover, we sympathize with the argument that authoritarian governments

---

5. The ReadMe software generates proportions of documents within categories without coding individual documents (Hopkins and King 2010, 230).

6. King et al.'s fig. 9 does not report a sample size. We base this estimate on conversions from the standard error sizes in the graph.

are less concerned about criticism than conventional wisdom would suggest. Yet, the claim that China's leaders are indifferent toward criticism, as long as it lacks CAP, is questionable. Instead, we propose that China's leaders selectively condition public discourse by explicitly demarcating areas over which criticism is acceptable. To be sure, there is little reason to doubt the importance of CAP. Indeed, it may very well be the paramount consideration. However, this does not mean that authoritarian regimes are unperturbed by unsolicited criticism that paints them in a bad light.

## REFERENCES

Chen, Jidong, and Yiqing Xu. 2017. "Why Do Authoritarian Regimes Allow Citizens to Voice Opinions Publicly?" *Journal of Politics* 79 (3): 792–803.

Dimitrov, Martin K. 2017. "The Political Logic of Media Control in China." *Problems of Post-Communism* 64 (3–4): 121–27.

*Economist*. 2017. "Cambridge University Press Battles Censorship in China." *Economist*, August 24. https://goo.gl/qoi7WP.

Esarey, Ashley, and Qiang Xiao. 2008. "Political Expression in the Chinese Blogosphere: Below the Radar." *Asian Survey* 48 (5): 752–72.

Fitzpatrick, Sheila. 1996. "Supplicants and Citizens: Public Letter-Writing in Soviet Russia in the 1930s." *Slavic Review* 55 (1): 78–105.

Geddes, Barbara, and John Zaller. 1989. "Sources of Popular Support for Authoritarian Regimes." *American Journal of Political Science* 33 (2): 319–47.

Gueorguiev, Dimitar. 2014. "Retrofitting Communism: Consultative Autocracy in China." PhD thesis, University of California, San Diego.

He, Baogang, and Mark E. Warren. 2011. "Authoritarian Deliberation: The Deliberative Turn in Chinese Political Development." *Perspectives on Politics* 9 (2): 269–89.

Hopkins, Daniel J., and Gary King. 2010. "A Method of Automated Nonparametric Content Analysis for Social Science." *American Journal of Political Science* 54 (1): 229–47.

Horsley, Jamie P. 2009. "Public Participation in the People's Republic: Developing a More Participatory Governance Model in China." Unpublished manuscript (September), Yale China Law Center.

Inkeles, Alex, and Kent Geiger. 1953. "Critical Letters to the Editors of the Soviet Press: Social Characteristics and Interrelations of Critics and the Criticized." *American Sociological Review* 18 (1): 12–22.

King, Gary, Jennifer Pan, and Margaret Roberts. 2013. "How Censorship in China Allows Government Criticism but Silences Collective Expression." *American Political Science Review* 107 (2): 326–43.

King, Gary, Jennifer Pan, and Margaret Roberts. 2014. "Reverse-Engineering Censorship in China: Randomized Experimentation and Participant Observation." *Science* 345 (6199): 1251722.

Koetse, Manya. 2017. "Storm of Criticism over Proposal to Lower China's Legal Marriage Age to 18." What's on Weibo: Reporting Social Trends in China, March 4. https://goo.gl/w6yJrq (accessed June 30, 2017).

McDonell, Stephen. 2017. "China Congress: How Authorities Censor Your Thoughts." BBC News, October 16. https://goo.gl/QAJi9f.

Miller, Blake. 2017. "The Limits of Commercialized Censorship in China." Working paper.

Stromseth, Jonathan, Edmund Malesky, and Dimitar Gueorguiev. 2017. *China's Governance Puzzle: Enabling Transparency and Participation in a Single-Party State*. Cambridge: Cambridge University Press.

Truex, Rory. 2014. "Consultative Authoritarianism and Its Limits." *Comparative Political Studies* 14 (3): 1–33.