Replication and extension of How Chinese Officials Use the Internet to Construct their Public Image by Jennifer Pan

Alexander Klueber

4/7/2020

Contents

1 Abstract

Pan shows that the emphasis on Chinese local government websites on either the competence or benevolence of county executives depends on where they are in the political tenure cycle. Early tenure county executives project images of benevolence by emphasizing their attentiveness and concern toward citizens. Late tenure executives project images of competence by highlighting their achievements. These findings shift the nature of debates concerning the role of the Internet in authoritarian regimes from a focus on regime-society interactions to an examination of dynamics among regime insiders. I was largely able to replicate the statistical models that she uses to suggest the before mentioned relationships. My own extension confirms that this is the most likely explanation for the observed effect by introducing a series of models that could support alternative explanations for the observed effect (e.g. cultural differences among regions, gender differences, etc.) and comparing the explanatory power of these models through the leave one out method. In addition I could validate the randomness of the sample selected to draw the underlying conclusions from a geographical perspective through a repeated sampling simulation and the construction of confidence intervals and their comparison with the observed provinces in the sample of 100 and the sample of 48 that was ultimately used to construct the models.

2 Introduction

I am interested to explore whether other explanations than the signalling function within authoritarian regimes may plausibly explain the alterations in competence / benevolence patterns described in the paper.

I will explore these alternatives by comparing the explanatory power of the variable categories employed in the paper and extending them with a new category (culture which will include the macro-region and the county type). Alternative hypothesis therefore are: 1. The benevolence/competence patterns may be explained by regional cultural variations 2. The benevolence/competence patterns may be explained by the resources at disposal to the official 3. The benevolence/competence patterns may be explained by internal peer preferences

4. The benevolence/competence patterns may be explained by characteristics of the prefecture 5. The benevolence/competence patterns may be explained by the individual abilities of the county officials 6. The benevolence/competence patterns may be explained by the immediate career success of county officials

Proving that it is in fact not tenure, but one of the other factors that have nothing to do with the career path of the individuals within the party, would weaken that connection. This comparison between models seems relevant as in the original paper the variables that are statistically significant vary between the models. In the model with the most controls, for competence "mayor education levels" and "whether a county party secretary is in first year of office" are statistically significant, for benevolence no variables are statistically significant. Between the regressions the number of observations included also varies, which initially makes a comparison between the models and the variables in the models difficult.

This endevour is constrained by the data available in the dataverse of the replication paper. Only the pre-selected sample of 100 contains all variables relevant for the analysis on a county level. Therefore a sub-division of that data by provinces will leave us with very small sample sizes. Other examples are that the sample only contains 4 female mayors, 1 person with education level 6, 4 people with education level 5, 3 people with education levle 2 and 2 people with education level 0.

Building on this, the other part of my extension is around validating the sample selection process. Initially, it seems surprising that the author has sub-selected 100 of the 2,787 counties with website, and we immediately dismiss another 29 counties in our modelling because of data inavailability and subsequently another ~23 as our regressions get more ambitious (include more variables). The absence of any website from Tibet is furthermore conspicuous.

3 Literature review

The replication paper aims to contribute to the larger discusison around role of the Internet in authoritarian regimes, especially in China. It is part of a larger shift in embracing the complexity of the role of the internet within these regimes. It moved from a prevailing assumption of the Internet's inherent democratic nature, and its ascribed power to undermine authoritarian regimes to a much more nuanced view that emphasises the utilitization of the internet in relation to how it is playing out.(Kalathil and Boas 2003)

While transparency laws are implemented by central authorities with the intention to obtain more information about the performance of local officials, they turn out to be political instruments for self-promotion among regime insiders. Research shows that the desired transparency does not come to frutition because grievances submitted through these online forums are systematically concealed from upper-level authorities when they implicate lower-tier officials or associates connected to lower-tier officialsthrough patronage ties. Information manipulation occurs primarily through omission of wrongdoingrather than censorship or falsification, suggesting that even in the digital age, in a highly determined andcapable regime where reports of corruption are actively and publicly voiced, monitoring the behavior ofregime agents remains a challenge.(Pan and Chen 2018)

There is however some disagreement whether creating transparency on local government performance is the primary function of these laws to begin with, or whether they are primarily intended as subtle instruments of online social control through information delivery, agenda setting, and containment of public dissent.(**???**)

The emphasis of local officials on the communication among insiders are also reflected by field experiments testing the responsiveness of local officials. Tattling to upper levels of government made county governments considerably more responsive to citizen's demands.(Chen, Pan, and Xu 2016) This challenges to some degree earlier findings that the capacity of the central state to monitor and control lower level agents has increased

in China as it suggests that some of the efforts to do so have altered the nature of the interaction rather than shifting the power relationships within it.(Edin 2003)

This isn't necessarily something negative in of itself, as empirical findings suggest that China uses personnel control to induce desirable outcomes, especially economic gains.(Li and Zhou 2005) This alludes to an associated discussion around the relevance of image building in authoritarian regimes and the missapropriation of resources in the service of that. China is mentioned as an example of that where the political arrangements have greacted the incentive and the opportunities for irresponsible behaviour among state agents to do so.(Cai 2004)

4 Extension

4.1 Comparison of alternative regression models in explaining competence and benevolence

One challenge with the regressions is that they are actually based on different underlying data. This is because of the data availability in the sample of 100 counties and then the prediction file. As we employ the various regressions, the number of observations used to fit these regressions decreases from 71 to 48. They are therefore fit based on different underlying data.

In a first step I therefore harmonize the regressions by basing them on the same number of observations throughout - the 48 observations that have all data available. I then compare whether the new regressions yield similar results as the regressions in the paper.

4.1.1 Competence

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Mon, Apr 13, 2020 - 16:55:03

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Beginning Tenure	4,000	0.170	0.027	0.075	0.152	0.188	0.279
End Tenure	4,000	0.007	0.049	−0.148	−0.026	0.040	0.182
mayor_last	4,000	0.136	0.059	−0.069	0.100	0.174	0.365
sigma	4,000	0.145	0.015	0.105	0.134	0.154	0.231

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Mon, Apr 13, 2020 - 16:55:03

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Beginning Tenure	4,000	0.188	0.055	0.010	0.152	0.225	0.398
End Tenure	4,000	−0.007	0.056	−0.238	−0.044	0.030	0.228
mayor_last	4,000	0.130	0.063	−0.151	0.089	0.172	0.393
X2009_gdppc_cny	4,000	0.00000	0.00000	−0.00000	0.00000	0.00000	0.00000
X2010_illiterateprop	4,000	−0.004	0.004	−0.020	−0.006	−0.001	0.011
itemploy	4,000	−0.0001	0.0001	−0.0003	−0.0001	−0.0001	0.0001
linksall	4,000	0.00000	0.00000	−0.00000	0.00000	0.00000	0.00000
county_mediaexp	4,000	−0.00000	0.00000	−0.00000	−0.00000	0.00000	0.00000
sigma	4,000	0.145	0.017	0.098	0.133	0.155	0.217

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Mon, Apr 13, 2020 - 16:55:03

3

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Beginning Tenure	4,000	0.156	0.059	−0.081	0.118	0.194	0.360
End Tenure	4,000	−0.013	0.057	−0.273	−0.051	0.025	0.213
mayor_last	4,000	0.127	0.069	−0.101	0.080	0.173	0.382
X2009_gdppc_cny	4,000	0.00000	0.00000	−0.00000	0.00000	0.00000	0.00000
X2010_illiterateprop	4,000	−0.003	0.004	−0.018	−0.006	−0.0005	0.014
itemploy	4,000	−0.0001	0.0001	−0.0003	−0.0001	−0.00005	0.0001
linksall	4,000	0.00000	0.00000	−0.00000	0.00000	0.00000	0.00000
county_mediaexp	4,000	−0.00000	0.00000	−0.00000	−0.00000	0.00000	0.00000
sec_first	4,000	0.125	0.060	−0.071	0.086	0.163	0.329
sec_last	4,000	0.040	0.058	−0.163	0.003	0.078	0.228
sigma	4,000	0.140	0.016	0.098	0.129	0.149	0.221

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Beginning Tenure	4,000	0.253	0.152	−0.360	0.151	0.356	0.884
End Tenure	4,000	0.014	0.066	−0.221	−0.030	0.057	0.281
mayor_last	4,000	0.169	0.076	−0.106	0.120	0.220	0.491
X2009_gdppc_cny	4,000	0.00000	0.00000	−0.00000	0.00000	0.00000	0.00001
X2010_illiterateprop	4,000	−0.002	0.005	−0.021	−0.005	0.001	0.017
itemploy	4,000	−0.0001	0.0001	−0.0003	−0.0001	−0.00001	0.0002
linksall	4,000	0.00000	0.00000	−0.00000	0.00000	0.00000	0.00000
county_mediaexp	4,000	0.00000	0.00000	−0.00000	−0.00000	0.00000	0.00000
sec_first	4,000	0.111	0.062	−0.122	0.070	0.153	0.379
sec_last	4,000	0.049	0.060	−0.220	0.010	0.088	0.267
pref_ps_first	4,000	0.080	0.078	−0.205	0.027	0.131	0.393
pref_ps_last	4,000	0.082	0.073	−0.171	0.033	0.128	0.450
pref_ps_edulevel	4,000	−0.032	0.029	−0.139	−0.052	−0.012	0.106
pref_2010_gdppc	4,000	−0.00000	0.00000	−0.00001	−0.00000	−0.00000	0.00000
sigma	4,000	0.140	0.017	0.097	0.128	0.151	0.219

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Beginning Tenure	4,000	0.425	0.303	−0.826	0.228	0.624	1.700
End Tenure	4,000	−0.045	0.068	−0.321	−0.089	−0.001	0.238
mayor_last	4,000	0.154	0.072	−0.150	0.107	0.204	0.396
X2009_gdppc_cny	4,000	0.00000	0.00000	−0.00000	0.000	0.00000	0.00001
X2010_illiterateprop	4,000	−0.001	0.005	−0.017	−0.004	0.002	0.015
itemploy	4,000	0.00001	0.0001	−0.0003	−0.00004	0.0001	0.0003
linksall	4,000	0.00000	0.00000	−0.00000	−0.00000	0.00000	0.00000
county_mediaexp	4,000	−0.00000	0.00000	−0.00000	−0.00000	0.00000	0.00000
sec_first	4,000	0.152	0.063	−0.079	0.110	0.193	0.483
sec_last	4,000	0.029	0.057	−0.177	−0.009	0.065	0.251
pref_ps_first	4,000	0.071	0.077	−0.249	0.019	0.122	0.400
pref_ps_last	4,000	0.107	0.070	−0.138	0.060	0.154	0.353
pref_ps_edulevel	4,000	−0.049	0.030	−0.169	−0.069	−0.030	0.063
pref_2010_gdppc	4,000	−0.00000	0.00000	−0.00000	−0.00000	0.00000	0.00000
mayor_age	4,000	−0.006	0.006	−0.031	−0.010	−0.001	0.019
mayor_genderM	4,000	−0.077	0.101	−0.391	−0.145	−0.009	0.310
mayor_edulevel	4,000	0.063	0.031	−0.061	0.043	0.083	0.189
sigma	4,000	0.136	0.018	0.087	0.123	0.146	0.232

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Beginning Tenure	4,000	0.459	0.310	−0.725	0.261	0.666	1.755
End Tenure	4,000	−0.052	0.072	−0.313	−0.100	−0.005	0.223
mayor_last	4,000	0.170	0.093	−0.197	0.110	0.232	0.524
X2009_gdppc_cny	4,000	0.00000	0.00000	−0.00000	−0.000	0.00000	0.00000
X2010_illiterateprop	4,000	−0.001	0.005	−0.020	−0.004	0.002	0.019
itemploy	4,000	0.00001	0.0001	−0.0003	−0.00004	0.0001	0.0003
linksall	4,000	0.00000	0.00000	−0.00000	−0.00000	0.00000	0.00000
county_mediaexp	4,000	−0.00000	0.00000	−0.00000	−0.00000	−0.000	0.00000
sec_first	4,000	0.150	0.068	−0.101	0.106	0.194	0.388
sec_last	4,000	0.029	0.059	−0.186	−0.009	0.068	0.289
pref_ps_first	4,000	0.072	0.083	−0.232	0.015	0.128	0.461
pref_ps_last	4,000	0.107	0.073	−0.171	0.059	0.156	0.379
pref_ps_edulevel	4,000	−0.051	0.030	−0.170	−0.072	−0.031	0.070
pref_2010_gdppc	4,000	−0.00000	0.00000	−0.00001	−0.00000	0.00000	0.00000
mayor_age	4,000	−0.006	0.006	−0.033	−0.011	−0.002	0.018
mayor_genderM	4,000	−0.075	0.103	−0.462	−0.142	−0.008	0.308
mayor_edulevel	4,000	0.067	0.031	−0.049	0.047	0.088	0.177
mayor_promote	4,000	−0.022	0.072	−0.326	−0.070	0.025	0.273
sigma	4,000	0.138	0.018	0.088	0.125	0.149	0.225

4.1.2 Benevolence

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Beginning Tenure	4,000	0.183	0.032	0.081	0.162	0.205	0.306
End Tenure	4,000	0.090	0.055	−0.150	0.052	0.127	0.337
mayor_last	4,000	−0.018	0.071	−0.285	−0.065	0.029	0.245
sigma	4,000	0.169	0.018	0.116	0.156	0.180	0.262

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Mon, Apr 13, 2020 - 16:55:32

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Beginning Tenure	4,000	0.154	0.063	−0.072	0.111	0.194	0.399
End Tenure	4,000	0.147	0.065	−0.102	0.103	0.192	0.369
mayor_last	4,000	0.005	0.071	−0.255	−0.043	0.055	0.245
X2009_gdppc_cny	4,000	−0.00000	0.00000	−0.00001	−0.00000	−0.00000	0.00000
X2010_illiterateprop	4,000	0.003	0.005	−0.014	0.0002	0.007	0.022
itemploy	4,000	0.00001	0.0001	−0.0002	−0.00004	0.0001	0.0003
linksall	4,000	0.00000	0.00000	−0.00000	0.00000	0.00000	0.00000
county_mediaexp	4,000	0.00000	0.00000	−0.00000	−0.00000	0.00000	0.00000
sigma	4,000	0.168	0.019	0.115	0.154	0.179	0.265

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Mon, Apr 13, 2020 - 16:55:33

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Beginning Tenure	4,000	0.147	0.069	−0.098	0.102	0.193	0.433
End Tenure	4,000	0.144	0.069	−0.119	0.097	0.191	0.375
mayor_last	4,000	0.010	0.082	−0.283	−0.044	0.064	0.322
X2009_gdppc_cny	4,000	−0.00000	0.00000	−0.00001	−0.00000	−0.00000	0.00000
X2010_illiterateprop	4,000	0.003	0.005	−0.016	0.0003	0.007	0.020
itemploy	4,000	0.00001	0.0001	−0.0003	−0.00004	0.0001	0.0003
linksall	4,000	0.00000	0.00000	−0.00000	0.00000	0.00000	0.00000
county_mediaexp	4,000	0.00000	0.00000	−0.00000	−0.00000	0.00000	0.00000
sec_first	4,000	0.043	0.074	−0.274	−0.006	0.092	0.290
sec_last	4,000	0.005	0.067	−0.246	−0.038	0.051	0.260
sigma	4,000	0.171	0.019	0.123	0.157	0.182	0.258

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Mon, Apr 13, 2020 - 16:55:33

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Beginning Tenure	4,000	0.213	0.191	−0.596	0.083	0.342	0.923
End Tenure	4,000	0.142	0.085	−0.168	0.086	0.200	0.447
mayor_last	4,000	0.018	0.094	−0.376	−0.045	0.083	0.360
X2009_gdppc_cny	4,000	−0.00000	0.00000	−0.00001	−0.00000	−0.00000	0.00000
X2010_illiterateprop	4,000	0.003	0.006	−0.019	−0.001	0.008	0.030
itemploy	4,000	0.00002	0.0001	−0.0004	−0.00004	0.0001	0.0004
linksall	4,000	0.00000	0.00000	−0.00000	0.00000	0.00000	0.00000
county_mediaexp	4,000	0.00000	0.00000	−0.00000	−0.00000	0.00000	0.00000
sec_first	4,000	0.044	0.080	−0.240	−0.009	0.095	0.359
sec_last	4,000	0.010	0.076	−0.246	−0.041	0.058	0.310
pref_ps_first	4,000	0.028	0.102	−0.400	−0.039	0.096	0.386
pref_ps_last	4,000	−0.010	0.089	−0.310	−0.069	0.049	0.323
pref_ps_edulevel	4,000	−0.019	0.037	−0.148	−0.044	0.006	0.127
pref_2010_gdppc	4,000	−0.00000	0.00000	−0.00001	−0.00000	0.00000	0.00001
sigma	4,000	0.180	0.022	0.125	0.164	0.193	0.306

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Beginning Tenure	4,000	0.194	0.411	−1.504	−0.068	0.460	1.806
End Tenure	4,000	0.146	0.094	−0.194	0.085	0.206	0.502
mayor_last	4,000	0.026	0.099	−0.343	−0.042	0.091	0.398
X2009_gdppc_cny	4,000	−0.00000	0.00000	−0.00001	−0.00000	−0.00000	0.00000
X2010_illiterateprop	4,000	0.003	0.007	−0.019	−0.001	0.008	0.033
itemploy	4,000	0.00002	0.0001	−0.0004	−0.0001	0.0001	0.0004
linksall	4,000	0.00000	0.00000	−0.00000	0.00000	0.00000	0.00000
county_mediaexp	4,000	0.00000	0.00000	−0.00000	−0.00000	0.00000	0.00000
sec_first	4,000	0.039	0.085	−0.257	−0.019	0.094	0.466
sec_last	4,000	0.008	0.080	−0.282	−0.045	0.062	0.332
pref_ps_first	4,000	0.013	0.105	−0.361	−0.056	0.084	0.541
pref_ps_last	4,000	−0.021	0.095	−0.333	−0.084	0.043	0.374
pref_ps_edulevel	4,000	−0.016	0.040	−0.146	−0.043	0.011	0.140
pref_2010_gdppc	4,000	0.00000	0.00000	−0.00001	−0.00000	0.00000	0.00001
mayor_age	4,000	−0.002	0.009	−0.037	−0.007	0.004	0.030
mayor_genderM	4,000	0.110	0.132	−0.421	0.020	0.199	0.568
mayor_edulevel	4,000	−0.006	0.041	−0.175	−0.033	0.021	0.162
sigma	4,000	0.186	0.024	0.125	0.169	0.200	0.346

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Beginning Tenure	4,000	0.078	0.396	−1.410	−0.190	0.336	1.650
End Tenure	4,000	0.169	0.092	−0.195	0.109	0.229	0.496
mayor_last	4,000	−0.062	0.119	−0.475	−0.143	0.017	0.420
X2009_gdppc_cny	4,000	−0.00000	0.00000	−0.00001	−0.00000	−0.00000	0.00000
X2010_illiterateprop	4,000	0.004	0.006	−0.019	−0.0003	0.008	0.028
itemploy	4,000	0.00001	0.0001	−0.0004	−0.0001	0.0001	0.0005
linksall	4,000	0.00000	0.00000	−0.00000	0.00000	0.00000	0.00000
county_mediaexp	4,000	0.00000	0.00000	−0.00000	−0.00000	0.00000	0.00000
sec_first	4,000	0.056	0.085	−0.246	0.0004	0.112	0.393
sec_last	4,000	−0.0001	0.078	−0.293	−0.052	0.052	0.280
pref_ps_first	4,000	−0.006	0.103	−0.363	−0.074	0.062	0.418
pref_ps_last	4,000	−0.021	0.093	−0.370	−0.083	0.040	0.378
pref_ps_edulevel	4,000	−0.013	0.039	−0.154	−0.039	0.013	0.150
pref_2010_gdppc	4,000	0.00000	0.00000	−0.00001	−0.00000	0.00000	0.00001
mayor_age	4,000	0.0003	0.009	−0.032	−0.005	0.006	0.034
mayor_genderM	4,000	0.115	0.132	−0.394	0.029	0.204	0.604
mayor_edulevel	4,000	−0.013	0.041	−0.167	−0.039	0.014	0.137
mayor_promote	4,000	0.122	0.092	−0.219	0.061	0.184	0.459
sigma	4,000	0.183	0.023	0.126	0.166	0.196	0.378

In a second step I go through the various regressions seeking to understand whether any of the tested variable classes (resources, peers, prefecture, ability, career path) explain the observed phenomena better than the ones around tenure. I do so by creating a series of new regressions that include only the variables in the respective variable classes. In addition I introduce a new series of classes: culture that includes the variables macro-region and county type.

I then compare all the available models with the leave-one-out method to see which one of these is best suited to explain the observed phenomena.

```
##           elpd_diff se_diff
## linear_11  0.0       0.0
## linear_1  -0.1       1.6
## linear_10 -0.2       2.8
## linear_8  -0.2       3.7
## linear_12 -2.4       2.1
## linear_7  -2.6       3.2
## linear_2  -2.8       2.5
## linear_3  -3.1       4.4
## linear_9  -3.7       2.4
## linear_5  -6.6       4.9
## linear_4  -7.5       4.5
## linear_6  -8.1       5.0
```

```
##           elpd_diff se_diff
## linear.1    0.0       0.0
## linear.11  -0.5       2.1
## linear.12  -1.4       2.0
## linear.8   -1.6       1.7
## linear.9   -2.4       2.4
## linear.10  -2.6       2.1
## linear.2   -2.6       2.1
## linear.7   -3.9       2.4
```

```
## linear.3   -4.8        2.3
## linear.4  -10.0        2.8
## linear.6  -12.9        3.0
## linear.5  -13.2        2.9
```

The comparison suggests that the best models for how competent an official is described as, are the tenure based model employed in the paper and whether officials were promoted in the two ensuing years. The later is a post-treatment variable. Rather than delviering additional insights on what the factors around an official are that determine the website content, it suggests the relative importance of the website content in determining whether an individual is promoted. In combination this supports the hypothesis of the author that the websites have an important signalling function within the Chinese state apparatus to determine who gets promoted.

4.2   Validating geographic split of sample

Subsequently I am simulating repeated sampling. I do so by creating a function that allows me to draw 100 and 48 random samples from the underlying countywebsites (countyweb). I repeat this step 1000 times, counting the number of counties from each province. I then compare that count with the count of counties in the paper samples (the sample of 100 and the 48 counties we actually end up constructing a model with).

4.2.1   Random sample of 100

Distribution for frequency of counties per province in sample o

The geographic distribution of the sample utilized in the paper is surprising bas
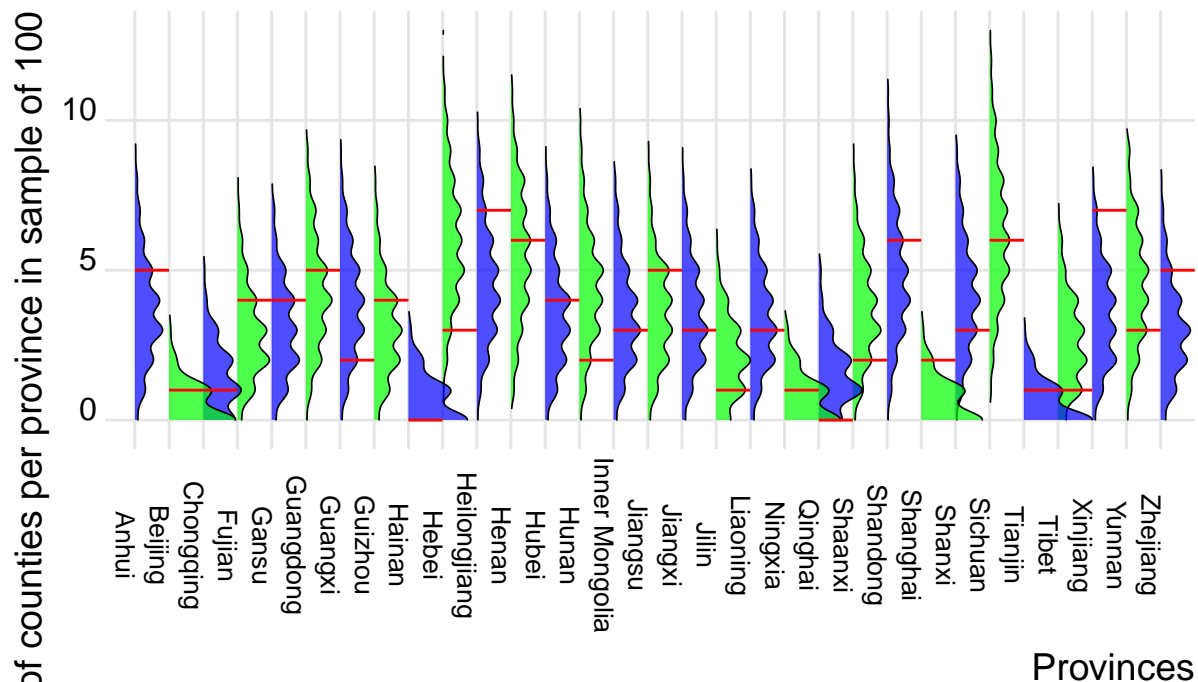
ta from paper: How Chinese Officials Use the Internet to Construct their Public Image

Extension Table 1: Frequency of provinces in sample of 100
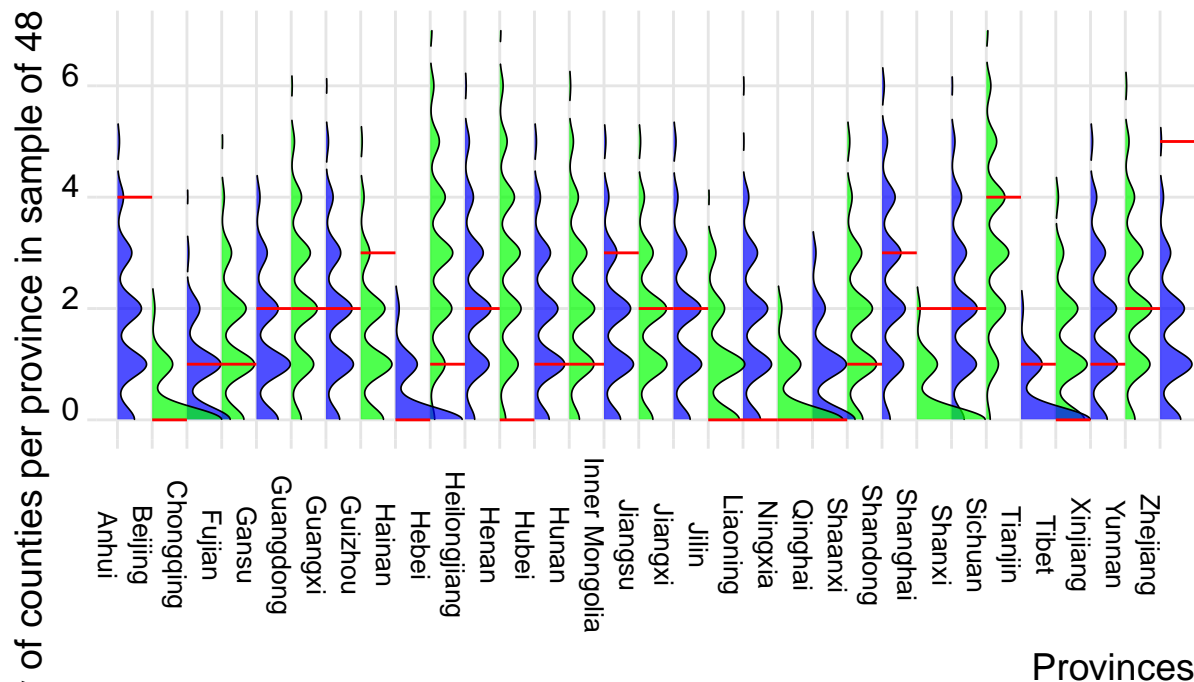Based on simulation of 1000 samples of 100

Provinces	2.5 Percentile	97.5 Percentile	Frequency in paper sample	Outside CI
Anhui	1.000	8.000	5	No
Beijing	0.000	2.000	1	No
Chongqing	0.000	4.000	1	No
Fujian	0.000	6.000	4	No
Gansu	0.000	7.000	4	No
Guangdong	1.000	8.000	5	No
Guangxi	1.000	8.000	2	No
Guizhou	0.000	7.000	4	No
Hainan	0.000	2.000	0	No
Hebei	2.000	11.000	3	No
Heilongjiang	1.000	9.000	7	No
Henan	1.000	10.000	6	No
Hubei	1.000	8.000	4	No
Hunan	1.000	9.000	2	No
Inner Mongolia	0.000	7.000	3	No
Jiangsu	1.000	8.000	5	No
Jiangxi	0.000	7.000	3	No
Jilin	0.000	5.000	1	No
Liaoning	0.000	7.000	3	No
Ningxia	0.000	3.000	1	No
Qinghai	0.000	4.000	0	No
Shaanxi	1.000	7.025	2	No
Shandong	1.000	10.000	6	No
Shanghai	0.000	2.000	2	No
Shanxi	1.000	8.000	3	No
Sichuan	2.000	11.000	6	No
Tianjin	0.000	2.000	1	No
Tibet	0.000	6.000	1	No
Xinjiang	0.975	7.000	7	No
Yunnan	1.000	9.000	3	No
Zhejiang	0.000	7.000	5	No

Repeated simulated sampling allows us to conclude that the sample of 100 in the paper is random and thereby representative in terms of geographic sampling. This is because the number of counties from a province in no case is outside the 95% CI interval that we constructed. The graph shows that in some counties, s.a. Heilongjiang or Hennan the county occurences are rather on the margins of what we would expect to see.

4.2.2 Sample of 48 for modelling

Distribution for frequency of counties per province in sample of

The geographic distribution of the sample utilized in the paper is surprising based

Data from paper: How Chinese Officials Use the Internet to Construct their Public Image

Extension Table 1: Frequency of provinces in sample of 48
Based on simulation of 1000 samples of 48

Provinces	2.5 Percentile	97.5 Percentile	Frequency in paper sample	Outside CI	On lower CI boundary
Anhui	0	4.000	4	No	No
Beijing	0	2.000	0	No	Yes
Chongqing	0	3.000	1	No	No
Fujian	0	4.000	1	No	No
Gansu	0	4.000	2	No	No
Guangdong	0	5.000	2	No	No
Guangxi	0	5.000	2	No	No
Guizhou	0	4.000	3	No	No
Hainan	0	2.000	0	No	Yes
Hebei	0	6.000	1	No	No
Heilongjiang	0	5.000	2	No	No
Henan	0	6.000	0	No	Yes
Hubei	0	5.000	1	No	No
Hunan	0	5.000	1	No	No
Inner Mongolia	0	4.025	3	No	No
Jiangsu	0	5.000	2	No	No
Jiangxi	0	5.000	2	No	No
Jilin	0	3.000	0	No	Yes
Liaoning	0	4.000	0	No	Yes
Ningxia	0	2.000	0	No	Yes

11

Qinghai	0	3.000	0	No	Yes
Shaanxi	0	5.000	1	No	No
Shandong	0	6.000	3	No	No
Shanghai	0	2.000	2	No	No
Shanxi	0	5.000	2	No	No
Sichuan	0	6.000	4	No	No
Tianjin	0	2.000	1	No	No
Tibet	0	4.000	0	No	Yes
Xinjiang	0	4.000	1	No	No
Yunnan	0	5.000	2	No	No
Zhejiang	0	4.000	5	Yes	No

Repeated simulated sampling allows us to conclude that the sample of 48 in the paper is likely random and thereby representative in terms of geographic sampling. This is because the number of counties from a province in only one case (Zhejiang) is outside the 95% CI interval that we constructed. The graph shows that there is a surprising amount of states at the lower boundary of 0. This seems plausible due to the small size of the sample (Beijing, Hainan, Henan, Jilin, Liaoning, Ningxia, Qinghai and Tibet). There seems to be no regional pattern among these states (3 East, 2 Central, 3 West).

References

Cai, Yongshun. 2004. "Irresponsible State: Local Cadres and Image-Building in China." *Journal of Communist Studies and Transition Politics* 20 (4): 20–41. https://doi.org/10.1080/1352327042000306039.

Chen, Jidong, Jennifer Pan, and Yiqing Xu. 2016. "Sources of Authoritarian Responsiveness: A Field Experiment in China." *American Journal of Political Science* 60 (2): 383–400. http://www.jstor.org/stable/24877628.

Edin, Maria. 2003. "State Capacity and Local Agent Control in China: CCP Cadre Management from a Township Perspective." *The China Quarterly*, no. 173: 35–52. http://www.jstor.org/stable/20058957.

Kalathil, Shanthi, and Taylor Boas. 2003. "Open Networks, Closed Regimes: The Impact of the Internet on Authoritarian Rule (Table of Contents)." *First Monday* 8 (1). https://doi.org/10.5210/fm.v8i1.1024.

Li, Hongbin, and Li-An Zhou. 2005. "Political Turnover and Economic Performance: The Incentive Role of Personnel Control in China." *Journal of Public Economics* 89 (9): 1743–62. https://doi.org/https://doi.org/10.1016/j.jpubeco.2004.06.009.

Pan, Jennifer, and Kaping Chen. 2018. *Concealing Corruption: How Chinese Officials Distort Upward Reporting of Online Grievances.* American Political Science Review 112.