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 (2017) 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. I was largely able to replicate these results. My extension confirms that this is the most likely explanation for the observed effect by comparing the statistical explanatory power of alternative models (e.g. cultural differences among regions, gender differences, etc.) through the leave one out method. In addition I validate the geographical randomness of the sample through simulations of repeated sampling and the construction of confidence intervals. They corraborate the findings of the paper by confirming the geographic randomness of the sample.

2 Introduction

The paper contributes to a larger discussion around the use of government websites in authroitarian regimes. Emplyoing the example of China it suggests that instead of enhancing transparency (as intended by the central authorities) these websites serve as vehicles of self-promotion and communication between regime insiders. To conclude this, Pan employes a random sample of 100 from 1.92 million county-level government web pages, classifies the messaging on individual sites (informative, competency or benevolence signalling) and regresses their respective frequency on a series of variables, always including their position in the tenure cycle. She concludes that the tenure of the county government officials and their subsequent promotion (a post-treatment variable) were most directive. This suggests the relevance of these websites in promotion decisions and the importance of tenure on the content being signalled.

To replicate this paper, I used R (R Core Team 2019). Original data and code of the replication paper are available in the Harvard dataverse.[1] The code for the extension is available at my repo.[2]

In addition to replicating the resulty, 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 individually (not layering) and extending them with the new category culture (includes the macro-region and the county type). Alternative hypothesis conclusively 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. 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.

I find that regressions including the tenure cycle are indeed those best able to explain the website content. The post-treatment promotion variable corrobates the importance of signalling through these websites further.

[1] https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/2HTWSU
[2] https://github.com/Alex1005-stack/Gov_1006_final_project

Furthermore, rebasing the sample on a consistent set of observations does not alter the conclusions we can draw. Lastly, the sample that constitutes the basis of this paper is constituted at random geographically.

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. The discussion moved from the 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 role of the internet in relation to how it is being employed by governments and people.(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 often turn out to be political instruments for self-promotion among regime insiders (as this paper shows). Research indicates further that the desired transparency does not come to frutition because grievances submitted through these websites are systematically concealed from upper-level authorities when they implicate lower-tier officials or associates connected to lower-tier officials through patronage ties. Information manipulation occurs primarily through omission of wrongdoing rather than censorship or falsification, suggesting that even in the digital age, in a highly determined and capable regime where reports of corruption are actively and publicly voiced, remains challenged to monitor the behavior of regime agents.(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 corraborated 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 shifted 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 missapropriation of resources in the service of individual image building. China is an example where the political arrangements have creacted the incentive and the opportunities for irresponsible behaviour among state agents to do so.(Cai 2004)

4 Paper Review

The paper analyses how local government officials use the websites they are required to put up due to the central government's transparency requirements, to engineer their public image. It concludes that the internet becomes a tool for self-promotion in authoritarian regimes. The analysis demonstrates that the websites tend to highlight the competence (achievements) or the benevolence (concern for citizens) of county executives which change based on the tenure of their cycle. The former is more important to officials that are later in their cycle, while the latter is more important to those earlier in the cycle.

The analysis does so by focusing on a random sample of 100 Chinese counties and their websites. They represent 29 Chinese provinces. Based on this sample they estimate the proportion of posts for competence and benevolence by year in office, where group refers to the year, with 95% confidence intervals in lwr and upr. Generated using Hopkins and King (2010) ReadMe software. The authors then conduct a linear regression that regresses a mayor first year or last year binary indicator on the number of competence metnions on the websites by county which they extend gradually for additional variables in alternative models. The same is then done for benevolence. These variable include: county resources and environmental factors, incentives of county party secretaries, incentives of prefecture party secretary and other prefecture-level characteristics, county executive's ability, age, and gender, post-treatment variable of whether the county executive was promoted. The authors conclude from the regression that When a county executive is in the last year of office, an additional 15% or so of website content is on average dedicated to claims of competence. Similarly they conclude that When a county executive is in the first year of office, an additional 10% to 15% of website content is on average dedicated to claims of benevolence. Both results are statistically significant at the 0.05 level.

5 Replication

5.1 Section 3.1: Website Content

5.1.1 Figure 1

I have not replicated the figure 1, a map of the various provinces in China and the availability of county government websites in them. This is because this was a manual step in the construction of the paper. I have replicated the data that was used as a basis for that map.

5.2 Section 4.1: Topics

5.2.1 Table 1

I have generated the underlying data for the table. I have not transferred it into the same table form as outlined in the paper.

5.3 Section 5.3: Predictive Inference

5.3.1 Table 3: Regression Results: Competence

Replicated the table, except positioning the constant at the bottom of the table

5.3.2 Table 4: Regression Results: Benevolence

Same as above

6 Extension

6.1 Comparison of alternative regression models in explaining competence and benevolence

One challenge with the regressions is that they are based on different underlying data. This is because of the data availability in the sample of 100 counties and the prediction file. As we employ the various regressions, the number of observations in our sample 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. There might only be a change in the values among the first four regressions (as (5) and (6) already are based on the 48 counties in the original replication paper). The subsequent tables represent the results side by side. The first column being the replication paper result, the second the result based on the 48 counties.

6.1.1 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 20, 2020 - 16:46:51

	Dependent variable: Competence							
Observations	71	48	70	48	70	48	68	48
Beginning Tenure	0.04	0.01	0.05	−0.01	0.05	−0.01	0.06	0.02
	(0.04)	(0.05)	(0.05)	(0.06)	(0.05)	(0.06)	(0.06)	(0.07)
End Tenure	0.15***	0.14**	0.14**	0.13**	0.13**	0.13*	0.15**	0.18**
	(0.05)	(0.06)	(0.05)	(0.06)	(0.06)	(0.07)	(0.06)	(0.07)
X2009_gdppc_cny			−0.0000	0.0000	−0.0000	0.0000	−0.0000	0.0000
			(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
X2010_illiterateprop			−0.004	−0.003	−0.005	−0.003	−0.004	−0.002
			(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)
itemploy			−0.0001*	−0.0001	−0.0001*	−0.0001	−0.0001	−0.0001
			(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
linksall			0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
			(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
county_mediaexp			0.0000	−0.0000	0.0000	−0.0000	0.0000	0.0000
			(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
sec_first					0.03	0.13**	0.03	0.11*
					(0.06)	(0.06)	(0.06)	(0.06)
sec_last					0.01	0.04	0.02	0.05
					(0.05)	(0.06)	(0.05)	(0.06)
pref_ps_first							−0.02	0.08
							(0.06)	(0.08)
pref_ps_last							0.06	0.08
							(0.07)	(0.07)
pref_ps_edulevel							−0.02	−0.03
							(0.02)	(0.03)
pref_2010_gdppc							−0.0000	−0.0000
							(0.0000)	(0.0000)
Constant	0.16***	0.17***	0.19***	0.19***	0.19***	0.16***	0.25*	0.25*
	(0.03)	(0.03)	(0.05)	(0.05)	(0.05)	(0.06)	(0.12)	(0.15)
Resource controls	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Peer controls	No	No	No	No	Yes	Yes	Yes	Yes
Prefecture controls	No	No	No	No	No	No	Yes	Yes
Ability controls	No	No	No	No	No	No	No	No
Career paths controls	No	No	No	No	No	No	No	No
Note:								*p<0.1; **p<0.05; ***p<0.01

6.1.2 Benevolence

% 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 20, 2020 - 16:46:52

	Dependent variable: Benevolence							
Observations	71	48	70	48	70	48	68	48
Beginning Tenure	0.07	0.09	0.10**	0.15**	0.10*	0.15**	0.13**	0.15*
	(0.05)	(0.06)	(0.05)	(0.06)	(0.05)	(0.07)	(0.06)	(0.08)
End Tenure	−0.003	−0.02	0.03	0.01	0.04	0.01	0.04	0.02
	(0.06)	(0.07)	(0.06)	(0.07)	(0.06)	(0.08)	(0.07)	(0.10)
X2009_gdppc_cny			−0.0000	−0.0000	−0.0000	−0.0000	−0.0000	−0.0000
			(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
X2010_illiterateprop			0.003	0.004	0.004	0.004	0.01	0.004
			(0.004)	(0.005)	(0.004)	(0.005)	(0.01)	(0.01)
itemploy			0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
			(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
linksall			0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
			(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
county_mediaexp			−0.0000	0.0000	−0.0000	0.0000	−0.0000	0.0000
			(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
sec_first					0.01	0.04	−0.01	0.04
					(0.06)	(0.07)	(0.07)	(0.08)
sec_last					−0.02	0.01	−0.03	0.01
					(0.05)	(0.07)	(0.06)	(0.07)
pref_ps_first							0.05	0.03
							(0.06)	(0.10)
pref_ps_last							0.01	−0.01
							(0.07)	(0.09)
pref_ps_edulevel							0.01	−0.02
							(0.03)	(0.04)
pref_2010_gdppc							0.0000	−0.0000
							(0.0000)	(0.0000)
Constant	0.19***	0.18***	0.15***	0.15**	0.15***	0.14**	0.08	0.21
	(0.03)	(0.03)	(0.05)	(0.06)	(0.05)	(0.07)	(0.13)	(0.19)
Resource controls	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Peer controls	No	No	No	No	Yes	Yes	Yes	Yes
Prefecture controls	No	No	No	No	No	No	Yes	Yes
Ability controls	No	No	No	No	No	No	No	No
Career paths controls	No	No	No	No	No	No	No	No

Note: *p<0.1; **p<0.05; ***p<0.01

A noteworthy difference is that in regards to competence the coeffcient for beginning tenure turns negative right away when based on the more limited 48 sample. It is worth to consider whether the change in the underlying sample is the most likely explanation for that shift (as this coefficient also is in no case statistically significant).

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.

6.1.3 Competence

```
##              elpd_diff se_diff
## linear_1_gl   0.0       0.0
## linear_11_gl -0.3       1.5
## linear_10_gl -0.4       3.3
## linear_8_gl  -0.7       4.1
## linear_12_gl -2.4       2.8
## linear_7_gl  -2.8       3.9
## linear_2_gl  -2.9       2.2
## linear_3_gl  -3.5       4.3
## linear_9_gl  -4.1       3.2
## linear_5_gl  -6.6       4.8
## linear_4_gl  -7.6       4.4
## linear_6_gl  -7.9       5.0
```

6.1.4 Benevolence

```
##              elpd_diff se_diff
## linear.1_gl    0.0       0.0
## linear.11_gl  -0.5       2.1
## linear.12_gl  -1.4       2.0
## linear.8_gl   -1.4       1.8
## linear.9_gl   -2.3       2.4
## linear.10_gl  -2.8       2.2
## linear.2_gl   -2.9       2.1
## linear.7_gl   -4.1       2.4
## linear.3_gl   -5.1       2.3
## linear.4_gl  -10.1       2.7
## linear.6_gl  -13.0       3.1
## linear.5_gl  -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.

Please see graphs and tablesprint the detailed results of only the 3 most relevant models for each regression (please see Recplication code and graphs and table section for all details).

```
## stan_glm
##  family:       gaussian [identity]
##  formula:      comp ~ mayor_first + mayor_last
##  observations: 48
##  predictors:   3
## ------
##             Median MAD_SD
## (Intercept) 0.2    0.0
## mayor_first 0.0    0.0
## mayor_last  0.1    0.1
```

```
##
## Auxiliary parameter(s):
##       Median MAD_SD
## sigma 0.1    0.0
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg

## stan_glm
##  family:       gaussian [identity]
##  formula:      comp ~ mayor_promote
##  observations: 48
##  predictors:   2
## ------
##               Median MAD_SD
## (Intercept)   0.2    0.0
## mayor_promote 0.1    0.0
##
## Auxiliary parameter(s):
##       Median MAD_SD
## sigma 0.1    0.0
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg

## stan_glm
##  family:       gaussian [identity]
##  formula:      comp ~ mayor_age + mayor_gender + mayor_edulevel
##  observations: 48
##  predictors:   4
## ------
##                Median MAD_SD
## (Intercept)     0.4    0.2
## mayor_age       0.0    0.0
## mayor_genderM  -0.1    0.1
## mayor_edulevel  0.0    0.0
##
## Auxiliary parameter(s):
##       Median MAD_SD
## sigma 0.1    0.0
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg

## stan_glm
##  family:       gaussian [identity]
##  formula:      benev ~ mayor_first + mayor_last
##  observations: 48
##  predictors:   3
## ------
```

```
##             Median MAD_SD
## (Intercept) 0.2    0.0
## mayor_first 0.1    0.1
## mayor_last  0.0    0.1
##
## Auxiliary parameter(s):
##       Median MAD_SD
## sigma 0.2    0.0
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg
```

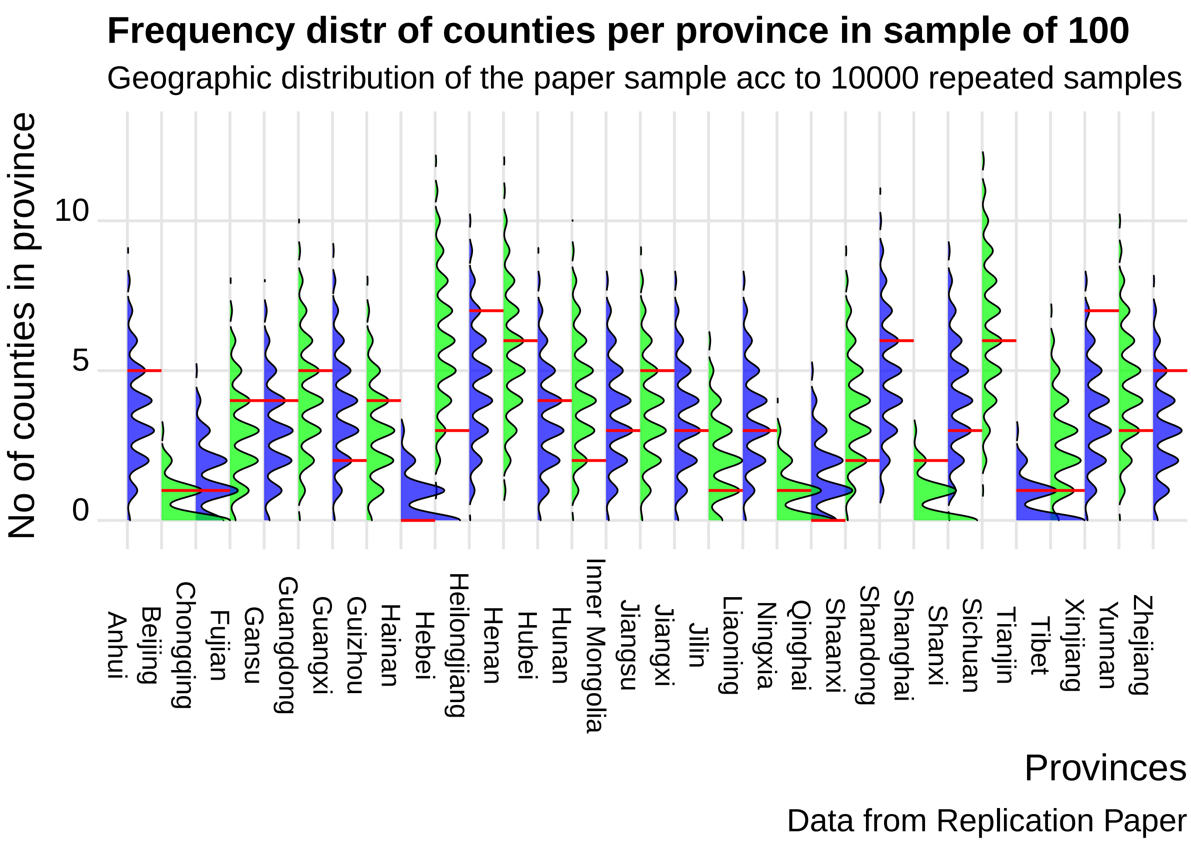
```
## stan_glm
##  family:       gaussian [identity]
##  formula:      benev ~ mayor_promote
##  observations: 48
##  predictors:   2
## ------
##               Median MAD_SD
## (Intercept)   0.2    0.0
## mayor_promote 0.0    0.1
##
## Auxiliary parameter(s):
##       Median MAD_SD
## sigma 0.2    0.0
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg
```

```
## stan_glm
##  family:       gaussian [identity]
##  formula:      benev ~ macroregion
##  observations: 48
##  predictors:   3
## ------
##                 Median MAD_SD
## (Intercept)      0.2    0.1
## macroregionEast  0.0    0.1
## macroregionWest -0.1    0.1
##
## Auxiliary parameter(s):
##       Median MAD_SD
## sigma 0.2    0.0
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg
```

6.2 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 10000 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).

6.2.1 Random sample of 100



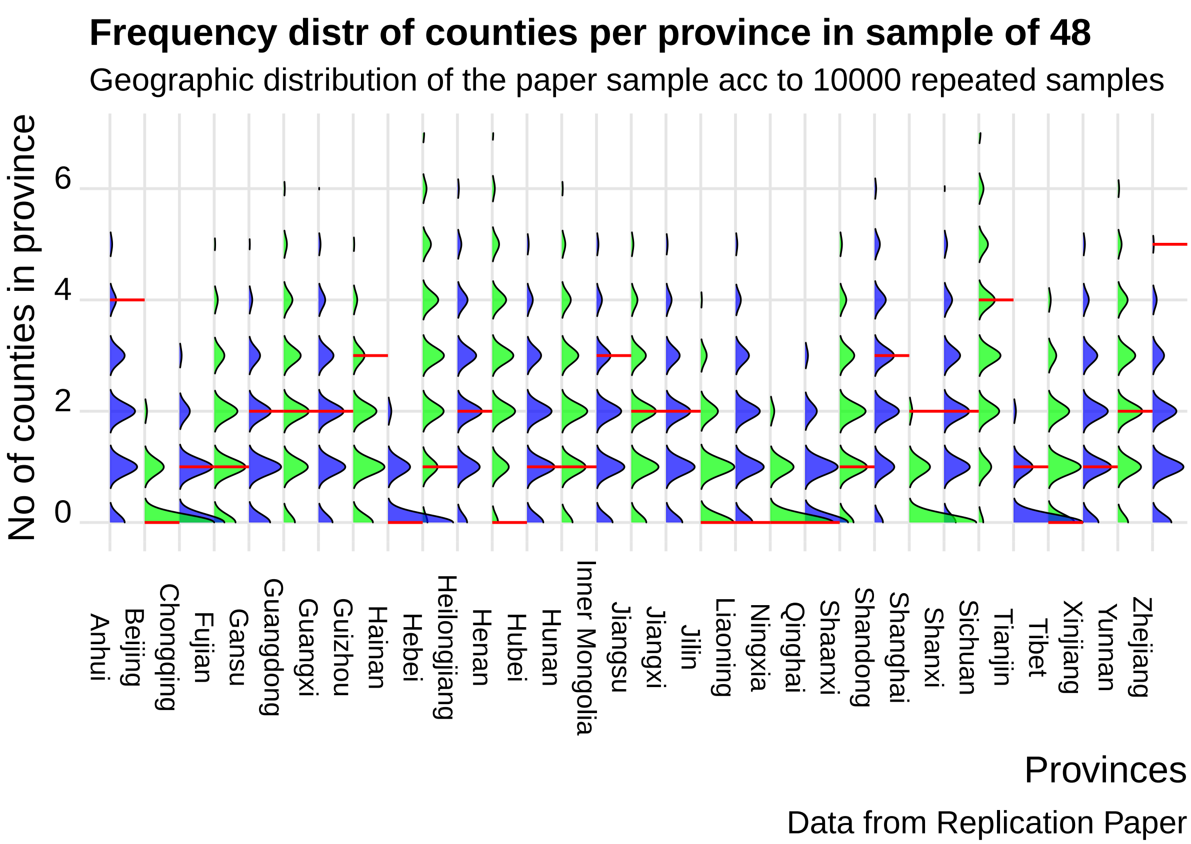
Extension Table 1: Frequency of provinces in sample of 100
Based on simulation of 10000 samples of 100

Provinces	2.5 Percentile	97.5 Percentile	Frequency in paper sample	Outside CI
Anhui	1	8	5	No
Beijing	0	2	1	No
Chongqing	0	4	1	No
Fujian	0	7	4	No
Gansu	0	7	4	No
Guangdong	1	8	5	No
Guangxi	1	8	2	No
Guizhou	0	7	4	No
Hainan	0	3	0	No
Hebei	2	11	3	No
Heilongjiang	1	9	7	No
Henan	2	10	6	No
Hubei	0	8	4	No
Hunan	1	8	2	No
Inner Mongolia	1	7	3	No
Jiangsu	1	8	5	No

Jiangxi	0	7	3	No
Jilin	0	5	1	No
Liaoning	0	7	3	No
Ningxia	0	3	1	No
Qinghai	0	4	0	No
Shaanxi	1	8	2	No
Shandong	1	9	6	No
Shanghai	0	2	2	No
Shanxi	1	8	3	No
Sichuan	2	11	6	No
Tianjin	0	2	1	No
Tibet	0	6	1	No
Xinjiang	1	8	7	No
Yunnan	1	9	3	No
Zhejiang	0	7	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.

6.2.2 Sample of 48 for modelling



Extension Table 2: Frequency of provinces in sample of 48
Based on simulation of 10000 samples of 48

Provinces	2.5 Percentile	97.5 Percentile	Frequency in paper sample	Outside CI	On lower CI boundary
Anhui	0	5	4	No	No
Beijing	0	2	0	No	Yes
Chongqing	0	3	1	No	No
Fujian	0	4	1	No	No
Gansu	0	4	2	No	No
Guangdong	0	5	2	No	No
Guangxi	0	5	2	No	No
Guizhou	0	4	3	No	No
Hainan	0	2	0	No	Yes
Hebei	0	6	1	No	No
Heilongjiang	0	5	2	No	No
Henan	0	6	0	No	Yes
Hubei	0	5	1	No	No
Hunan	0	5	1	No	No
Inner Mongolia	0	5	3	No	No
Jiangsu	0	5	2	No	No
Jiangxi	0	4	2	No	No
Jilin	0	3	0	No	Yes
Liaoning	0	5	0	No	Yes
Ningxia	0	2	0	No	Yes
Qinghai	0	3	0	No	Yes
Shaanxi	0	5	1	No	No
Shandong	0	6	3	No	No
Shanghai	0	2	2	No	No
Shanxi	0	5	2	No	No
Sichuan	0	7	4	No	No
Tianjin	0	2	1	No	No
Tibet	0	4	0	No	Yes
Xinjiang	0	5	1	No	No
Yunnan	0	5	2	No	No
Zhejiang	0	4	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).

7 Conclusion

The paper contributes to a larger discussion around the use of government websites in authroitarian regimes. Emplyoing the example of China it suggests that instead of enhancing transparency (as intended by the central authorities) these websites serve as vehicles of self-promotion and communication between regime insiders. To conclude this, Pan employes a random sample of 100 from 1.92 million county-level government web pages, classifies the messaging on individual sites (informative, competency or benevolence signalling) and regresses their respective frequency on a series of variables, always including their position in the tenure cycle. She concludes that the tenure of the county government officials and their subsequent promotion (a post-treatment variable) were most directive. This suggests the relevance of these websites in promotion decisions and the importance of tenure on the content being signalled.

To replicate this paper, I used R (R Core Team 2019). Original data and code of the replication paper are available in the Harvard dataverse.[3] The code for the extension is available at my repo.[4]

In addition to replicating the resulty, 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 individually (not layering) and extending them with the new category culture (includes the macro-region and the county type). Alternative hypothesis conclusively 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. 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).

I find that regressions including the tenure cycle are indeed those best able to explain the website content. The post-treatment promotion variable corrobates the importance of signalling through these websites further. Furthermore, rebasing the sample on a consistent set of observations does not alter the conclusions we can draw. Lastly, the sample that constitutes the basis of this paper is constituted at random geographically.

8 Tables and Figures

Below tables show the coefficients regressing a series of variables onto the number of competence mentions on the county government website. They compare a series of old regressions (using a diverging subset of the 100 county sample as the basis for the regression) with a series of new regressions (using the same 48 counties as the basis for all regressions and including a series of new regressions)

% 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 20, 2020 - 16:47:06

[3]https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/2HTWSU
[4]https://github.com/Alex1005-stack/Gov_1006_final_project

	Dependent variable: Competence							
Observations	71	48	70	48	70	48	68	48
Beginning Tenure	0.04	0.01	0.05	−0.01	0.05	−0.01	0.06	0.02
	(0.04)	(0.05)	(0.05)	(0.06)	(0.05)	(0.06)	(0.06)	(0.07)
End Tenure	0.15***	0.14**	0.14**	0.13**	0.13**	0.13*	0.15**	0.18**
	(0.05)	(0.06)	(0.05)	(0.06)	(0.06)	(0.07)	(0.06)	(0.07)
X2009_gdppc_cny			−0.0000	0.0000	−0.0000	0.0000	−0.0000	0.0000
			(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
X2010_illiterateprop			−0.004	−0.003	−0.005	−0.003	−0.004	−0.002
			(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)
itemploy			−0.0001*	−0.0001	−0.0001*	−0.0001	−0.0001	−0.0001
			(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
linksall			0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
			(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
county_mediaexp			0.0000	−0.0000	0.0000	−0.0000	0.0000	0.0000
			(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
sec_first					0.03	0.13**	0.03	0.11*
					(0.06)	(0.06)	(0.06)	(0.06)
sec_last					0.01	0.04	0.02	0.05
					(0.05)	(0.06)	(0.05)	(0.06)
pref_ps_first							−0.02	0.08
							(0.06)	(0.08)
pref_ps_last							0.06	0.08
							(0.07)	(0.07)
pref_ps_edulevel							−0.02	−0.03
							(0.02)	(0.03)
pref_2010_gdppc							−0.0000	−0.0000
							(0.0000)	(0.0000)
Constant	0.16***	0.17***	0.19***	0.19***	0.19***	0.16***	0.25*	0.25*
	(0.03)	(0.03)	(0.05)	(0.05)	(0.05)	(0.06)	(0.12)	(0.15)
Resource controls	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Peer controls	No	No	No	No	Yes	Yes	Yes	Yes
Prefecture controls	No	No	No	No	No	No	Yes	Yes
Ability controls	No	No	No	No	No	No	No	No
Career paths controls	No	No	No	No	No	No	No	No
Note:								*p<0.1; **p<0.05; ***p<0.01

Below tables show the coefficients regressing a series of variables onto the number of benevolence mentions on the county government website. They compare a series of old regressions (using a diverging subset of the 100 county sample as the basis for the regression) with a series of new regressions (using the same 48 counties as the basis for all regressions and including a series of new regressions)

% 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 20, 2020 - 16:47:07

	Dependent variable: Benevolence							
Observations	71	48	70	48	70	48	68	48
Beginning Tenure	0.07	0.09	0.10**	0.15**	0.10*	0.15**	0.13**	0.15*
	(0.05)	(0.06)	(0.05)	(0.06)	(0.05)	(0.07)	(0.06)	(0.08)
End Tenure	−0.003	−0.02	0.03	0.01	0.04	0.01	0.04	0.02
	(0.06)	(0.07)	(0.06)	(0.07)	(0.06)	(0.08)	(0.07)	(0.10)
X2009_gdppc_cny			−0.0000	−0.0000	−0.0000	−0.0000	−0.0000	−0.0000
			(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
X2010_illiterateprop			0.003	0.004	0.004	0.004	0.01	0.004
			(0.004)	(0.005)	(0.004)	(0.005)	(0.01)	(0.01)
itemploy			0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
			(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
linksall			0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
			(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
county_mediaexp			−0.0000	0.0000	−0.0000	0.0000	−0.0000	0.0000
			(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
sec_first					0.01	0.04	−0.01	0.04
					(0.06)	(0.07)	(0.07)	(0.08)
sec_last					−0.02	0.01	−0.03	0.01
					(0.05)	(0.07)	(0.06)	(0.07)
pref_ps_first							0.05	0.03
							(0.06)	(0.10)
pref_ps_last							0.01	−0.01
							(0.07)	(0.09)
pref_ps_edulevel							0.01	−0.02
							(0.03)	(0.04)
pref_2010_gdppc							0.0000	−0.0000
							(0.0000)	(0.0000)
Constant	0.19***	0.18***	0.15***	0.15**	0.15***	0.14**	0.08	0.21
	(0.03)	(0.03)	(0.05)	(0.06)	(0.05)	(0.07)	(0.13)	(0.19)
Resource controls	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Peer controls	No	No	No	No	Yes	Yes	Yes	Yes
Prefecture controls	No	No	No	No	No	No	Yes	Yes
Ability controls	No	No	No	No	No	No	No	No
Career paths controls	No	No	No	No	No	No	No	No

Note: *p<0.1; **p<0.05; ***p<0.01

This table compares and ranks them in order of their statistical explanatory power using the leave one out method.

Competence:

```
##              elpd_diff se_diff
## linear_1_gl   0.0       0.0
## linear_11_gl -0.3       1.5
## linear_10_gl -0.4       3.3
## linear_8_gl  -0.7       4.1
## linear_12_gl -2.4       2.8
## linear_7_gl  -2.8       3.9
## linear_2_gl  -2.9       2.2
## linear_3_gl  -3.5       4.3
## linear_9_gl  -4.1       3.2
## linear_5_gl  -6.6       4.8
```

```
## linear_4_gl  -7.6       4.4
## linear_6_gl  -7.9       5.0
```

Benevolence:

```
##              elpd_diff se_diff
## linear.1_gl    0.0       0.0
## linear.11_gl  -0.5       2.1
## linear.12_gl  -1.4       2.0
## linear.8_gl   -1.4       1.8
## linear.9_gl   -2.3       2.4
## linear.10_gl  -2.8       2.2
## linear.2_gl   -2.9       2.1
## linear.7_gl   -4.1       2.4
## linear.3_gl   -5.1       2.3
## linear.4_gl  -10.1       2.7
## linear.6_gl  -13.0       3.1
## linear.5_gl  -13.2       2.9
```

Printing all regressions and their results from extension

Competence

```
## stan_glm
##  family:       gaussian [identity]
##  formula:      comp ~ mayor_first + mayor_last
##  observations: 48
##  predictors:   3
## ------
##             Median MAD_SD
## (Intercept) 0.2    0.0
## mayor_first 0.0    0.0
## mayor_last  0.1    0.1
##
## Auxiliary parameter(s):
##       Median MAD_SD
## sigma 0.1    0.0
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg
```

```
## stan_glm
##  family:       gaussian [identity]
##  formula:      comp ~ mayor_first + mayor_last + X2009_gdppc_cny + X2010_illiterateprop +
##     itemploy + linksall + county_mediaexp
##  observations: 48
##  predictors:   8
## ------
##                     Median MAD_SD
## (Intercept)         0.2    0.1
## mayor_first         0.0    0.1
## mayor_last          0.1    0.1
```

```
## X2009_gdppc_cny      0.0    0.0
## X2010_illiterateprop 0.0    0.0
## itemploy             0.0    0.0
## linksall             0.0    0.0
## county_mediaexp      0.0    0.0
##
## Auxiliary parameter(s):
##       Median MAD_SD
## sigma 0.1    0.0
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg
```

```
## stan_glm
##  family:       gaussian [identity]
##  formula:      comp ~ mayor_first + mayor_last + X2009_gdppc_cny + X2010_illiterateprop +
##     itemploy + linksall + county_mediaexp + sec_first + sec_last
##  observations: 48
##  predictors:   10
## ------
##                      Median MAD_SD
## (Intercept)          0.2    0.1
## mayor_first          0.0    0.1
## mayor_last           0.1    0.1
## X2009_gdppc_cny      0.0    0.0
## X2010_illiterateprop 0.0    0.0
## itemploy             0.0    0.0
## linksall             0.0    0.0
## county_mediaexp      0.0    0.0
## sec_first            0.1    0.1
## sec_last             0.0    0.1
##
## Auxiliary parameter(s):
##       Median MAD_SD
## sigma 0.1    0.0
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg
```

```
## stan_glm
##  family:       gaussian [identity]
##  formula:      comp ~ mayor_first + mayor_last + X2009_gdppc_cny + X2010_illiterateprop +
##     itemploy + linksall + county_mediaexp + sec_first + sec_last +
##     pref_ps_first + pref_ps_last + pref_ps_edulevel + pref_2010_gdppc
##  observations: 48
##  predictors:   14
## ------
##                      Median MAD_SD
## (Intercept)          0.3    0.2
## mayor_first          0.0    0.1
## mayor_last           0.2    0.1
## X2009_gdppc_cny      0.0    0.0
```

```
## X2010_illiterateprop 0.0    0.0
## itemploy             0.0    0.0
## linksall             0.0    0.0
## county_mediaexp      0.0    0.0
## sec_first            0.1    0.1
## sec_last             0.0    0.1
## pref_ps_first        0.1    0.1
## pref_ps_last         0.1    0.1
## pref_ps_edulevel     0.0    0.0
## pref_2010_gdppc      0.0    0.0
##
## Auxiliary parameter(s):
##       Median MAD_SD
## sigma 0.1    0.0
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg
```

```
## stan_glm
##  family:       gaussian [identity]
##  formula:      comp ~ mayor_first + mayor_last + X2009_gdppc_cny + X2010_illiterateprop +
##     itemploy + linksall + county_mediaexp + sec_first + sec_last +
##     pref_ps_first + pref_ps_last + pref_ps_edulevel + pref_2010_gdppc +
##     mayor_age + mayor_gender + mayor_edulevel
##  observations: 48
##  predictors:   17
## ------
##                      Median MAD_SD
## (Intercept)           0.4    0.3
## mayor_first           0.0    0.1
## mayor_last            0.2    0.1
## X2009_gdppc_cny       0.0    0.0
## X2010_illiterateprop  0.0    0.0
## itemploy              0.0    0.0
## linksall              0.0    0.0
## county_mediaexp       0.0    0.0
## sec_first             0.2    0.1
## sec_last              0.0    0.1
## pref_ps_first         0.1    0.1
## pref_ps_last          0.1    0.1
## pref_ps_edulevel     -0.1    0.0
## pref_2010_gdppc       0.0    0.0
## mayor_age             0.0    0.0
## mayor_genderM        -0.1    0.1
## mayor_edulevel        0.1    0.0
##
## Auxiliary parameter(s):
##       Median MAD_SD
## sigma 0.1    0.0
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg
```

```
## stan_glm
##  family:       gaussian [identity]
##  formula:      comp ~ mayor_first + mayor_last + X2009_gdppc_cny + X2010_illiterateprop +
##     itemploy + linksall + county_mediaexp + sec_first + sec_last +
##     pref_ps_first + pref_ps_last + pref_ps_edulevel + pref_2010_gdppc +
##     mayor_age + mayor_gender + mayor_edulevel + mayor_promote
##  observations: 48
##  predictors:   18
## ------
##                      Median MAD_SD
## (Intercept)           0.5    0.3
## mayor_first          -0.1    0.1
## mayor_last            0.2    0.1
## X2009_gdppc_cny       0.0    0.0
## X2010_illiterateprop  0.0    0.0
## itemploy              0.0    0.0
## linksall              0.0    0.0
## county_mediaexp       0.0    0.0
## sec_first             0.1    0.1
## sec_last              0.0    0.1
## pref_ps_first         0.1    0.1
## pref_ps_last          0.1    0.1
## pref_ps_edulevel     -0.1    0.0
## pref_2010_gdppc       0.0    0.0
## mayor_age             0.0    0.0
## mayor_genderM        -0.1    0.1
## mayor_edulevel        0.1    0.0
## mayor_promote         0.0    0.1
##
## Auxiliary parameter(s):
##       Median MAD_SD
## sigma 0.1    0.0
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg
```

```
## stan_glm
##  family:       gaussian [identity]
##  formula:      comp ~ X2009_gdppc_cny + X2010_illiterateprop + itemploy + linksall +
##     county_mediaexp
##  observations: 48
##  predictors:   6
## ------
##                      Median MAD_SD
## (Intercept)          0.2    0.1
## X2009_gdppc_cny      0.0    0.0
## X2010_illiterateprop 0.0    0.0
## itemploy             0.0    0.0
## linksall             0.0    0.0
## county_mediaexp      0.0    0.0
##
## Auxiliary parameter(s):
##       Median MAD_SD
```

```
## sigma 0.1    0.0
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg

## stan_glm
##  family:       gaussian [identity]
##  formula:      comp ~ sec_first + sec_last
##  observations: 48
##  predictors:   3
## ------
##             Median MAD_SD
## (Intercept) 0.2    0.0
## sec_first   0.1    0.1
## sec_last    0.1    0.0
##
## Auxiliary parameter(s):
##       Median MAD_SD
## sigma 0.1    0.0
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg

## stan_glm
##  family:       gaussian [identity]
##  formula:      comp ~ pref_ps_first + pref_ps_last + pref_ps_edulevel + pref_2010_gdppc
##  observations: 48
##  predictors:   5
## ------
##                  Median MAD_SD
## (Intercept)      0.3    0.1
## pref_ps_first    0.1    0.1
## pref_ps_last     0.1    0.1
## pref_ps_edulevel 0.0    0.0
## pref_2010_gdppc  0.0    0.0
##
## Auxiliary parameter(s):
##       Median MAD_SD
## sigma 0.2    0.0
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg

## stan_glm
##  family:       gaussian [identity]
##  formula:      comp ~ mayor_age + mayor_gender + mayor_edulevel
##  observations: 48
##  predictors:   4
## ------
##                Median MAD_SD
```

```
## (Intercept)     0.4    0.2
## mayor_age       0.0    0.0
## mayor_genderM  -0.1    0.1
## mayor_edulevel  0.0    0.0
##
## Auxiliary parameter(s):
##       Median MAD_SD
## sigma 0.1    0.0
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg
```

```
## stan_glm
##  family:       gaussian [identity]
##  formula:      comp ~ mayor_promote
##  observations: 48
##  predictors:   2
## ------
##               Median MAD_SD
## (Intercept)   0.2    0.0
## mayor_promote 0.1    0.0
##
## Auxiliary parameter(s):
##       Median MAD_SD
## sigma 0.1    0.0
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg
```

```
## stan_glm
##  family:       gaussian [identity]
##  formula:      comp ~ macroregion
##  observations: 48
##  predictors:   3
## ------
##                 Median MAD_SD
## (Intercept)     0.2    0.0
## macroregionEast 0.0    0.1
## macroregionWest 0.0    0.1
##
## Auxiliary parameter(s):
##       Median MAD_SD
## sigma 0.2    0.0
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg
```

Benevolence

```
## stan_glm
```

```
##  family:       gaussian [identity]
##  formula:      benev ~ mayor_first + mayor_last
##  observations: 48
##  predictors:   3
## ------
##             Median MAD_SD
## (Intercept) 0.2    0.0   
## mayor_first 0.1    0.1   
## mayor_last  0.0    0.1   
##
## Auxiliary parameter(s):
##       Median MAD_SD
## sigma 0.2    0.0   
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg

## stan_glm
##  family:       gaussian [identity]
##  formula:      benev ~ mayor_first + mayor_last + X2009_gdppc_cny + X2010_illiterateprop + 
##     itemploy + linksall + county_mediaexp
##  observations: 48
##  predictors:   8
## ------
##                      Median MAD_SD
## (Intercept)          0.2    0.1   
## mayor_first          0.1    0.1   
## mayor_last           0.0    0.1   
## X2009_gdppc_cny      0.0    0.0   
## X2010_illiterateprop 0.0    0.0   
## itemploy             0.0    0.0   
## linksall             0.0    0.0   
## county_mediaexp      0.0    0.0   
##
## Auxiliary parameter(s):
##       Median MAD_SD
## sigma 0.2    0.0   
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg

## stan_glm
##  family:       gaussian [identity]
##  formula:      benev ~ mayor_first + mayor_last + X2009_gdppc_cny + X2010_illiterateprop + 
##     itemploy + linksall + county_mediaexp + sec_first + sec_last
##  observations: 48
##  predictors:   10
## ------
##                      Median MAD_SD
## (Intercept)          0.1    0.1   
## mayor_first          0.1    0.1   
## mayor_last           0.0    0.1   
```

```
## X2009_gdppc_cny      0.0    0.0
## X2010_illiterateprop 0.0    0.0
## itemploy             0.0    0.0
## linksall             0.0    0.0
## county_mediaexp      0.0    0.0
## sec_first            0.0    0.1
## sec_last             0.0    0.1
##
## Auxiliary parameter(s):
##       Median MAD_SD
## sigma 0.2    0.0
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg
```

```
## stan_glm
##  family:       gaussian [identity]
##  formula:      benev ~ mayor_first + mayor_last + X2009_gdppc_cny + X2010_illiterateprop +
##     itemploy + linksall + county_mediaexp + sec_first + sec_last +
##     pref_ps_first + pref_ps_last + pref_ps_edulevel + pref_2010_gdppc
##  observations: 48
##  predictors:   14
## ------
##                      Median MAD_SD
## (Intercept)          0.2    0.2
## mayor_first          0.1    0.1
## mayor_last           0.0    0.1
## X2009_gdppc_cny      0.0    0.0
## X2010_illiterateprop 0.0    0.0
## itemploy             0.0    0.0
## linksall             0.0    0.0
## county_mediaexp      0.0    0.0
## sec_first            0.0    0.1
## sec_last             0.0    0.1
## pref_ps_first        0.0    0.1
## pref_ps_last         0.0    0.1
## pref_ps_edulevel     0.0    0.0
## pref_2010_gdppc      0.0    0.0
##
## Auxiliary parameter(s):
##       Median MAD_SD
## sigma 0.2    0.0
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg
```

```
## stan_glm
##  family:       gaussian [identity]
##  formula:      benev ~ mayor_first + mayor_last + X2009_gdppc_cny + X2010_illiterateprop +
##     itemploy + linksall + county_mediaexp + sec_first + sec_last +
##     pref_ps_first + pref_ps_last + pref_ps_edulevel + pref_2010_gdppc +
##     mayor_age + mayor_gender + mayor_edulevel
```

```
##  observations: 48
##  predictors:   17
## ------
##                      Median MAD_SD
## (Intercept)          0.2    0.4   
## mayor_first          0.1    0.1   
## mayor_last           0.0    0.1   
## X2009_gdppc_cny      0.0    0.0   
## X2010_illiterateprop 0.0    0.0   
## itemploy             0.0    0.0   
## linksall             0.0    0.0   
## county_mediaexp      0.0    0.0   
## sec_first            0.0    0.1   
## sec_last             0.0    0.1   
## pref_ps_first        0.0    0.1   
## pref_ps_last         0.0    0.1   
## pref_ps_edulevel     0.0    0.0   
## pref_2010_gdppc      0.0    0.0   
## mayor_age            0.0    0.0   
## mayor_genderM        0.1    0.1   
## mayor_edulevel       0.0    0.0   
## 
## Auxiliary parameter(s):
##       Median MAD_SD
## sigma 0.2    0.0   
## 
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg
```

```
## stan_glm
##  family:       gaussian [identity]
##  formula:      benev ~ mayor_first + mayor_last + X2009_gdppc_cny + X2010_illiterateprop + 
##     itemploy + linksall + county_mediaexp + sec_first + sec_last + 
##     pref_ps_first + pref_ps_last + pref_ps_edulevel + pref_2010_gdppc + 
##     mayor_age + mayor_gender + mayor_edulevel + mayor_promote
##  observations: 48
##  predictors:   18
## ------
##                      Median MAD_SD
## (Intercept)           0.1    0.4  
## mayor_first           0.2    0.1  
## mayor_last           -0.1    0.1  
## X2009_gdppc_cny       0.0    0.0  
## X2010_illiterateprop  0.0    0.0  
## itemploy              0.0    0.0  
## linksall              0.0    0.0  
## county_mediaexp       0.0    0.0  
## sec_first             0.1    0.1  
## sec_last              0.0    0.1  
## pref_ps_first         0.0    0.1  
## pref_ps_last          0.0    0.1  
## pref_ps_edulevel      0.0    0.0  
## pref_2010_gdppc       0.0    0.0  
```

```
## mayor_age            0.0    0.0
## mayor_genderM        0.1    0.1
## mayor_edulevel       0.0    0.0
## mayor_promote        0.1    0.1
##
## Auxiliary parameter(s):
##       Median MAD_SD
## sigma 0.2    0.0
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg
```

```
## stan_glm
##  family:       gaussian [identity]
##  formula:      benev ~ X2009_gdppc_cny + X2010_illiterateprop + itemploy + linksall +
##     county_mediaexp
##  observations: 48
##  predictors:   6
## ------
##                      Median MAD_SD
## (Intercept)          0.2    0.1
## X2009_gdppc_cny      0.0    0.0
## X2010_illiterateprop 0.0    0.0
## itemploy             0.0    0.0
## linksall             0.0    0.0
## county_mediaexp      0.0    0.0
##
## Auxiliary parameter(s):
##       Median MAD_SD
## sigma 0.2    0.0
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg
```

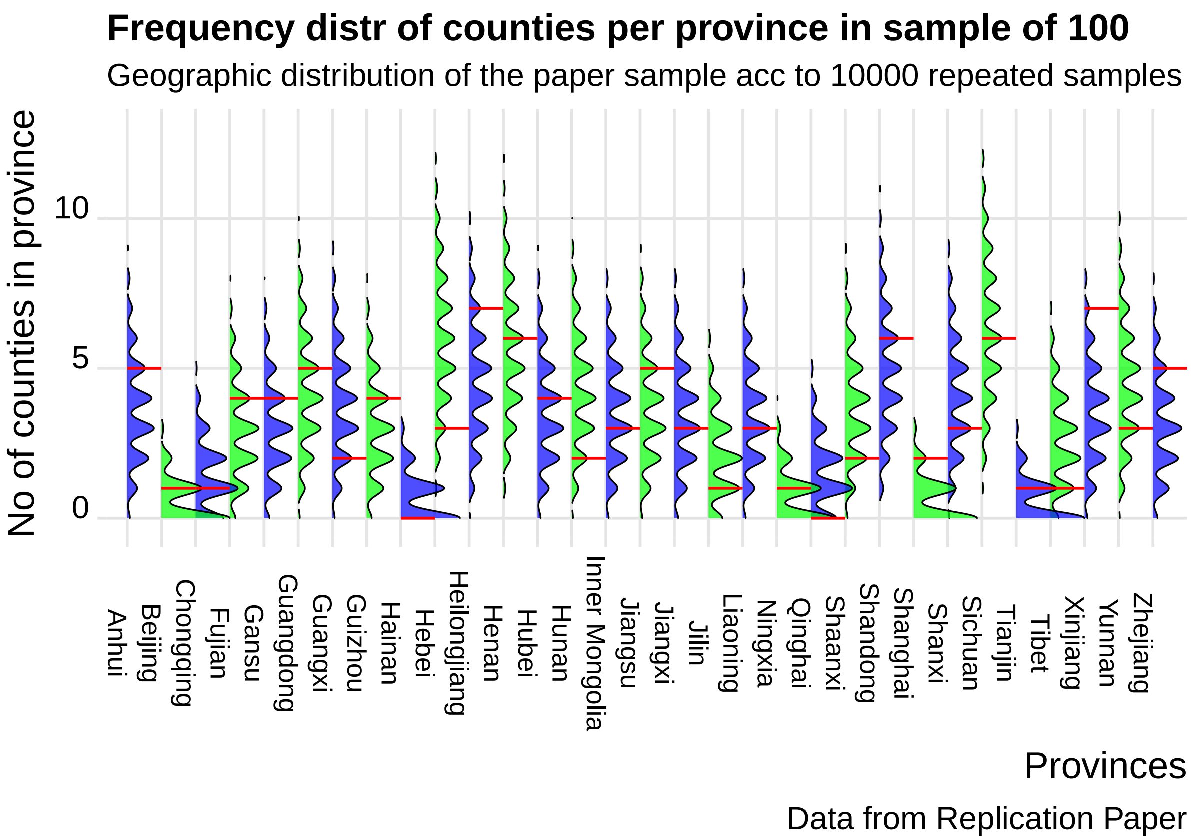
```
## stan_glm
##  family:       gaussian [identity]
##  formula:      benev ~ sec_first + sec_last
##  observations: 48
##  predictors:   3
## ------
##             Median MAD_SD
## (Intercept) 0.2    0.0
## sec_first   0.0    0.1
## sec_last    0.0    0.1
##
## Auxiliary parameter(s):
##       Median MAD_SD
## sigma 0.2    0.0
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg
```

```
## stan_glm
##  family:       gaussian [identity]
##  formula:      benev ~ pref_ps_first + pref_ps_last + pref_ps_edulevel + pref_2010_gdppc
##  observations: 48
##  predictors:   5
## ------
##                  Median MAD_SD
## (Intercept)       0.3    0.1
## pref_ps_first     0.0    0.1
## pref_ps_last     -0.1    0.1
## pref_ps_edulevel  0.0    0.0
## pref_2010_gdppc   0.0    0.0
##
## Auxiliary parameter(s):
##       Median MAD_SD
## sigma 0.2    0.0
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg

## stan_glm
##  family:       gaussian [identity]
##  formula:      benev ~ mayor_age + mayor_gender + mayor_edulevel
##  observations: 48
##  predictors:   4
## ------
##                Median MAD_SD
## (Intercept)    0.2    0.3
## mayor_age      0.0    0.0
## mayor_genderM  0.0    0.1
## mayor_edulevel 0.0    0.0
##
## Auxiliary parameter(s):
##       Median MAD_SD
## sigma 0.2    0.0
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg

## stan_glm
##  family:       gaussian [identity]
##  formula:      benev ~ mayor_promote
##  observations: 48
##  predictors:   2
## ------
##               Median MAD_SD
## (Intercept)   0.2    0.0
## mayor_promote 0.0    0.1
##
## Auxiliary parameter(s):
##       Median MAD_SD
## sigma 0.2    0.0
```

```
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg
```

```
## stan_glm
##  family:       gaussian [identity]
##  formula:      benev ~ macroregion
##  observations: 48
##  predictors:   3
## ------
##                 Median MAD_SD
## (Intercept)      0.2    0.1
## macroregionEast  0.0    0.1
## macroregionWest -0.1    0.1
##
## Auxiliary parameter(s):
##       Median MAD_SD
## sigma 0.2    0.0
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg
```

Drawing 10,000 repeated samples of 100 counties from the list of all county government websites, we graph the distribution of how many of the counties in the sample are in a particular province for all 31 provinces included in the original data set. The red lines indicate how many counties in a particular province were included the sample employed by the paper.



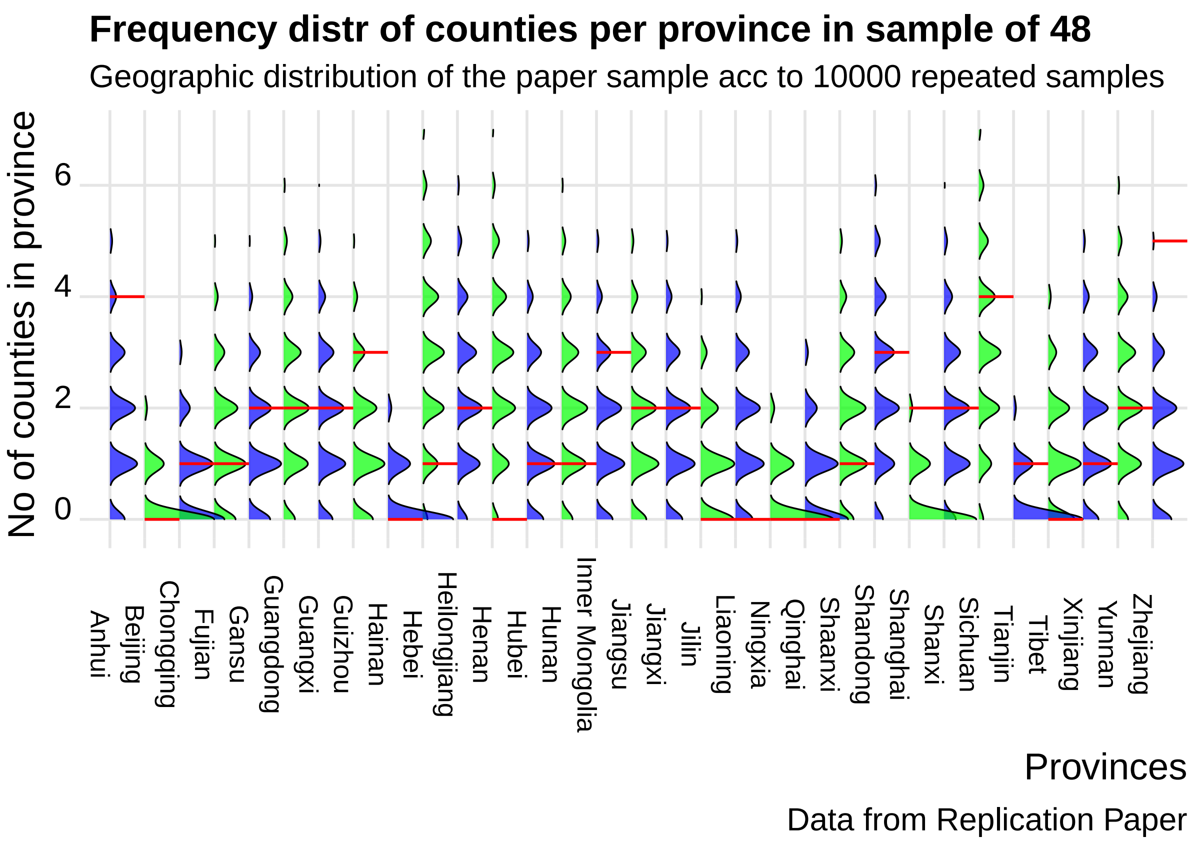
Based on the logic described above, the table then shows the limits of the 95% confidence intervals for that frequency and whether the observed frequency in the paper were outside these limits (0 counties).

Extension Table 1: Frequency of provinces in sample of 100
Based on simulation of 10000 samples of 100

Provinces	2.5 Percentile	97.5 Percentile	Frequency in paper sample	Outside CI
Anhui	1	8	5	No
Beijing	0	2	1	No
Chongqing	0	4	1	No
Fujian	0	7	4	No
Gansu	0	7	4	No
Guangdong	1	8	5	No
Guangxi	1	8	2	No
Guizhou	0	7	4	No
Hainan	0	3	0	No
Hebei	2	11	3	No
Heilongjiang	1	9	7	No
Henan	2	10	6	No
Hubei	0	8	4	No
Hunan	1	8	2	No
Inner Mongolia	1	7	3	No
Jiangsu	1	8	5	No
Jiangxi	0	7	3	No
Jilin	0	5	1	No
Liaoning	0	7	3	No

Ningxia	0	3	1	No
Qinghai	0	4	0	No
Shaanxi	1	8	2	No
Shandong	1	9	6	No
Shanghai	0	2	2	No
Shanxi	1	8	3	No
Sichuan	2	11	6	No
Tianjin	0	2	1	No
Tibet	0	6	1	No
Xinjiang	1	8	7	No
Yunnan	1	9	3	No
Zhejiang	0	7	5	No

Drawing 10,000 repeated samples of 48 counties from the list of all county government websites, we graph the distribution of how many of the counties in the sample are in a particular province for all 31 provinces included in the original data set. The red lines indicate how many counties in a particular province were included the sample employed by the paper.



Based on the logic described above, the table then shows the limits of the 95% confidence intervals for that frequency and whether the observed frequency in the paper were outside these limits (1 county) or on the lower boundary (8 counties).

Extension Table 2: Frequency of provinces in sample of 48
Based on simulation of 10000 samples of 48

Provinces	2.5 Percentile	97.5 Percentile	Frequency in paper sample	Outside CI	On lower CI boundary

Anhui	0	5	4	No	No
Beijing	0	2	0	No	Yes
Chongqing	0	3	1	No	No
Fujian	0	4	1	No	No
Gansu	0	4	2	No	No
Guangdong	0	5	2	No	No
Guangxi	0	5	2	No	No
Guizhou	0	4	3	No	No
Hainan	0	2	0	No	Yes
Hebei	0	6	1	No	No
Heilongjiang	0	5	2	No	No
Henan	0	6	0	No	Yes
Hubei	0	5	1	No	No
Hunan	0	5	1	No	No
Inner Mongolia	0	5	3	No	No
Jiangsu	0	5	2	No	No
Jiangxi	0	4	2	No	No
Jilin	0	3	0	No	Yes
Liaoning	0	5	0	No	Yes
Ningxia	0	2	0	No	Yes
Qinghai	0	3	0	No	Yes
Shaanxi	0	5	1	No	No
Shandong	0	6	3	No	No
Shanghai	0	2	2	No	No
Shanxi	0	5	2	No	No
Sichuan	0	7	4	No	No
Tianjin	0	2	1	No	No
Tibet	0	4	0	No	Yes
Xinjiang	0	5	1	No	No
Yunnan	0	5	2	No	No
Zhejiang	0	4	5	Yes	No

9 Appendixes

9.1 Section 3.1: Website Content

9.1.1 Figure 1: County Government Website Availability by Province

9.2 Section 4.1: Topics

9.2.1 Table 1: LDA topics and OGI Requirements

9.3 Section 5.1: Measuring Tenure

9.3.1 Table 2: Distribution of Year in office

9.4 Section 5.2: Descriptive Results

9.4.1 Figure 2: Proportion of web pages with content focused on competence by year in office...

9.4.2 Figure 3: Proportion of web pages with content focused on belevolence by year in office...

9.5 Section 5.3: Predictive Inference

9.5.1 Table 3: Regression Results: Competence

9.5.2 Table 4: Regression Results: Benevolence

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.

R Core Team. 2019. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.