

# From Internet to Social Safety Net: The Policy Consequences of Online Participation in China<sup>\*</sup>

Junyan Jiang<sup>†</sup>

Tianguang Meng<sup>‡</sup>

Qing Zhang<sup>§</sup>

## Abstract

Internet-based platforms are increasingly being used by governments around the world to facilitate public engagement with citizens. However, it remains an open question whether participation through these platforms can actually enable citizens to influence policies. We address this question by studying the patterns and consequences of online participation at a major electronic petition platform in China, a country with the world's largest Internet-using population. Content analysis of over 900,000 petitions reveals that a substantial share of them concern lower-class issues and are originated from less developed rural and suburban areas. Linking variations in petition volumes to an original dataset of government policy priorities, we further show that online participation led governments to place greater emphasis on social welfare policies and to increase the coverage of a key low-income assistance program. These results underscore the potential of online participation as an important mechanism to improve the quality of governance.

---

<sup>\*</sup>Earlier versions of the paper have been presented at APSA, SPSA, and the Young Economists Symposium at Yale. For helpful discussion and comments, we thank David Blei, Bruce Bueno de Mesquita, Anthony Fowler, Dan Hopkins, Holger Kern, Amit Khandelwal, Pierre Landry, Feng Li, Christian Goebel, W. Bentley MacLeod, Suresh Naidu, Ora John Reuter, Victor Shih, David Szakonyi, Eric Verhoogen, and Dali Yang, and seminar participants at University of Amsterdam, Columbia, Hong Kong Institute of Education, New York University, and UC San Diego. We are grateful to Manfred Elfstrom for generously sharing his labor protest data.

<sup>†</sup> Assistant Professor, Department of Government and Public Administration, the Chinese University of Hong Kong.

<sup>‡</sup> Associate Professor, Department of Political Science, Tsinghua University

<sup>§</sup> PhD Candidate, Department of Economics, Columbia University.

# 1 Introduction

Effective use of digital technologies is an essential part of modern governance. In the past two decades, national and local governments around the world have set up new Internet-based interfaces to interact with their citizens. In countries as diverse as Japan, United States, China, Colombia, and Ukraine, online participation (also known as e-participation) has become a common method for citizens to contact political authority and articulate their issues and grievances (United Nations, 2016, p.56). These developments have also triggered an intense debate among scholars about the nature and policy ramifications of online participation. While optimists view online participation as having the potential to expand civic engagement, enhance responsiveness of public institutions, and promote transparency and equity in policy making (Bimber, 1998; Castells, 1996), others remain skeptical about its ability to benefit underrepresented constituencies and its effectiveness in inducing broad policy changes (Chadwick & May, 2003; Lindner & Riehm, 2011).

In this study, we contribute new evidence to this debate by analyzing online participation in China. After almost two decades of rapid expansion of the Internet, China now hosts the largest Internet-using population in the world.<sup>1</sup> The Chinese government has been very active in creating various types of e-participation initiatives, including electronic petition platforms, online public consultation, and government-sponsored social media accounts.<sup>2</sup> However, the question remains whether these initiatives are just for window-dressing purposes, or whether they have genuine influence over government policies. To answer this question, our analysis focuses specifically on the *Local Leader Message Board* (LLMB), a major online petition forum that allows citizens to directly register complaints to party and government leaders in their localities on the Internet. Launched by China's central media since 2008, the LLMB exemplifies a new class of government-sponsored electronic participation platforms that are now being adopted worldwide (see Table A.1 for a list of similar institutions in other countries/regions). Compared to the more conventional

---

<sup>1</sup>According to China Internet Network Information Center (CNNIC), the number of Internet users in China is estimated to be over 710 million as of July 2016.

<sup>2</sup>As of 2016, China is ranked at 22 among the top 50 performers in e-participation (United Nations, 2016), a standing on par with Denmark and higher than Luxembourg (43), Portugal (50), and Iceland (50).

participatory institutions, this platform offers an unusually cheap and transparent way for citizens to communicate their issues and concerns to the authority. As a result, we hypothesize that it would be especially attractive to lower-class citizens, who tend to be discriminated against in the “normal” political process and are usually most sensitive to changes in participation costs. Based on a large body of research suggesting that lower-income groups tend to hold stronger preference for redistribution (Alesina & Ferrara, 2005; Meltzer & Richard, 1981), we further postulate that expanded online participation will shift government policies towards placing greater emphasis on social welfare and redistributive issues, as the voices of the poor become better heard.

To evaluate these hypotheses, we first examine the content of petitions filed on the LLMB platform. We manually code issues and user backgrounds for a random sample of petitions, and then use a non-parametric content analysis method developed by Hopkins and King (2010) to estimate their distributions within the entire body of over 900,000 petitions. Our analysis shows that a substantial share of the LLMB petitions concern personal problems that are most likely to be experienced by lower-class citizens, including employment, neighborhood environment, and land-taking compensation. Consistent with this, we also find that a sizable share of the petitions are originated from rural and suburban areas, which are inhabited primarily by individuals who belong to the lower strata of the Chinese society.<sup>3</sup>

We then investigate how expanded online participation affected substantive policy priorities of local governments. To construct a consistent measure of government policies that can be compared across time and space, we apply natural language processing techniques on over 4,400 Government Work Reports (GWRs), which are comprehensive policy blueprints published annually by the Chinese government at various levels, and estimate the relative proportions of different “topics” (i.e., clusters of words) as a measure for government policy priorities. Our baseline results show that cities that receive a larger number of online petitions in a year tend to devote significantly higher proportions of government reports in the following year to a topic on social welfare, which broadly

---

<sup>3</sup>In this article, we conceptualize the lower class in the Chinese context as including farmers and low-skill/migrant workers who usually live in rural areas or city outskirts (i.e., towns/suburban areas). This is consistent with the prevailing views about the social hierarchy in contemporary China (e.g., Goodman, 2014).

includes issues such as low-income subsidies, unemployment protection, and medical assistance. We also find that the policy effect appears to be mainly driven by petitions filed by *rural* residents concerning issues related to their *pocketbook conditions*. We subject these results to a number of robustness checks and adopt several different strategies to address the problem of endogeneity between online petition and government policy. Furthermore, leveraging data from the Minimum Living Standard Guarantee Scheme (*dibao*), a key low-income assistance program targeting the poor, we show that those changes in stated policy priorities are not merely empty talk, but are accompanied by substantive improvements in the coverage of the *dibao* program.

This paper contributes to the literature on e-governance. Researchers have investigated the nature and characteristics of different e-governance techniques (e.g., Chadwick & May, 2003; D. F. Norris & Reddick, 2012; Torres, Pina, & Acerete, 2006), the conditions under which certain e-governance techniques are adopted (e.g., Royo, Yetano, & Acerete, 2013; Wong & Welch, 2004), and the relationship between citizens' attitudes and their participatory experiences (e.g., Kim & Lee, 2012). However, there is still little systematic evidence on the critical question of whether online participation can have any substantive impact in government policies. Some research on advanced democracies suggests that e-government initiatives typically emphasize efficient service delivery over expanding participation opportunities for citizens (Chadwick & May, 2003; D. F. Norris & Reddick, 2012). Others analyze cross-national data and find that development in e-participation are not strongly associated with reduction in corruption perception or improvement in quality of government (e.g., Linde & Karlsson, 2013). Leveraging more fine-grained subnational data on online participation and government policies, our study offers evidence that e-participation can indeed pressure local governments to devote more attention to social welfare issues even in the absence of electoral accountability. These findings lend support to the view that the Internet and information technologies can be a tool of empowerment for citizens in the digital age (Dutton, 2009).

## 2 Background

### 2.1 Online Participation in China and the *Local Leader Message Board*

The rapid expansion of the Internet in China since the 1990s and especially during the first decade of the 21st century has brought profound changes to the way citizens and governments interact. The emergence of online forums, weblogs, and many other types of virtual communities have facilitated lively discussion of public affairs and provided citizens with new channels to articulate their problems and interests (Yang, 2009). At the same time, the government has also undertaken systematic efforts to expand its online presence. The word “E-government (电子政府)” first appeared in a major party policy document at the 16th Party Congress in 2002, and a series of regulations were subsequently promulgated to promote the development of central and local government websites, with a special emphasis on information availability and ease of access. As of 2008, virtually all local governments at or above the county-level had set up their own websites, and many used these websites regularly for disseminating information and delivering various types of services (Pan, 2017; Stromseth, Malesky, & Gueorguiev, 2017).

To understand the patterns and policy consequences of online participation in China, this article focuses on a major government-sponsored petition platform named the *Local Leader Message Board* (LLMB, 地方领导留言板). The LLMB was created in 2008 by the official website of China’s central media (人民网) with the goal of providing an integrated portal for citizens to contact leading government officials in their localities (for a snap shot of the interface, see Figure A.1 in Online Appendix). Ever since its launch, the platform has become a popular venue for citizens to report issues and register complaints to local authorities. As of August 2016, the LLMB has received over 900,000 petitions from all mainland provinces, making it the most heavily used platform in China today.<sup>4</sup>

---

<sup>4</sup>The LLMB is not the only online petition platform available in China. Inspired by the LLMB, many local governments later also developed similar platforms to receive and process petitions (for related studies, see Chen, Pan, and Xu (2016), Distelhorst and Hou (2017)). However, because of the concern for censorship and data availability issues (many local platforms do not make petition data publicly available), we only focus on the LLMB for this study. In Online Appendix D we compare LLMB petitions with petitions at a select set of local platforms on which we can get complete data. We find that the volumes of petitions from the two sources are highly correlated. This suggests that

LLMB hosts two message boards for each subnational unit at or above the county level (i.e., 33 provinces, 333 cities, and over 2,000 counties), one for the head of the party and the other for the head of the government. Participants can post their petitions to these message boards either anonymously or with a registered account. Registration requires only a functional cell phone number or an email address, and will allow the users to customize the names that appear on their petitions. After a petition is filed and before it is publicly posted, the website will conduct a quick check on the content. While some censorship does happen at this stage, our interviews with LLMB staff suggest that the criteria are much more lenient than those used for social media sites.<sup>5</sup> According to one member of the management team, petitions are typically only blocked if they contain vulgar languages or explicit comments/criticisms of top political leaders, and “99.9% of the petitions are displayed exactly the way they were written” (Personal interview, Beijing, December 2016).

Insofar as responsiveness is concerned, several features of the LLMB deserve special mentioning. First, in contrast to locally operated websites, in which local authorities have a direct control over the content to be displayed, the LLMB is operated by a central agency that typically has little direct interest in helping local authorities cover up their problems. This means that local governments will not be able to simply bury any undesirable issues through censorship. In addition, the website also employs several measures to increase the publicity of petitions and put pressure on local government to respond. For example, all petitions and government replies are publicly visible to all users once posted; this not only makes it easier for local residents with common grievances to find each other and organize, but can also help local leaders’ political superiors to learn about citizens’ opinions about their subordinates. Occasionally, the LLMB’s own news team will look for newsworthy materials from the petitions and conduct follow-up investigations that will be published on the central media’s website. All these features—the relative independence of the operator, the transparent design of the website, and the high publicity of the platform—may

---

petitions at LLMB can be seen as broadly representative of the general patterns of online petitions in China.

<sup>5</sup> This is in part because, unlike an online forum, LLMB users only interact with the government but not with each other; the lack of horizontal communication makes it less pressing to censor posts for the purpose of preventing collective actions.

give local officials an incentive to be attentive to citizen demands made on the platform.

### 3 Hypotheses

In this section, we develop hypotheses regarding (1) the type of citizens who would be drawn to the LLMB and (2) the potential policy consequences associated with online participation. First, we hypothesize that *lower-class citizens are more likely to use these online platforms than those from a better-off background*. The main function of the LLMB is facilitate citizens' *articulation of personal grievances* to local political leaders. According to the classical resource-mobilization theory of political participation, individualized interest articulation is a rather resource-demanding form of participation because there is no standardized menu for actions and the outcome depends largely on a citizen's own ability and initiatives (Marien, Hooghe, & Quintelier, 2010; Verba & Nie, 1972). This observation also holds in the context of China. Those who can afford living in major cities (i.e., closer to where the higher-level governments reside) or offering bribes to build personal connections with government officials are usually in a better position to influence government than those who lack such resources.<sup>6</sup> By drastically lowering the time and costs needed for long-distance communication, the LLMB can help improve rural residents' ability to contact higher-level governments located in distant urban centers. Its integrated and simplified interface also helps lower the knowledge barriers for navigating complex bureaucratic systems. Furthermore, the anonymous and impersonal nature of the website means that it would be much harder for officials to make differential treatment based on participants' political connections or other personal characteristics (Ward, Gibson, & Lusoli, 2003). In contrast to the upper-class citizens, whose background may give them an advantage in contacting government officials, anonymity is especially attractive to lower-class citizens, who are most likely to be discriminated against for

---

<sup>6</sup>Consistent with this view, existing research shows that the traditional off-line participation channels in China are disproportionately used by those of higher political and socioeconomic status. Guo (2007), for example, finds that members of the Chinese Communist Party (CCP) and various "mass organizations" have a higher propensity to utilize the official channels in political participation than do non-members. Tsai and Xu (Forthcoming) show that citizens with political connections are more likely to contact the authorities with complaints about public services.

their background.

It is also worth noting that, unlike many other developing countries, lower-class citizens in China do have the access to the necessary Internet *infrastructure* for effective online participation. Since the early 2000s, the government and state-owned telecommunication companies have made systematic efforts in building up Internet-related infrastructure (e.g., electricity, telephone lines, etc) in remote, rural areas (Oreglia, 2015). As of 2009, Internet access was available in over 90% of the administrative villages.<sup>7</sup> The recent advancement in mobile telecommunication technologies (i.e., 3G and 4G networks) has led to a further expansion of Internet users in rural areas. One estimate suggests that over 87% of the rural Internet users in China today access Internet primarily through smart phones (CNNIC, 2016). The wide availability of relatively inexpensive Internet access makes rural and lower-class citizens an influential group in China's cyberspace. Although direct evidence from political sphere remains limited, indications can be found from patterns of Internet usage in other areas: Research has shown, for example, that rural residents in China are highly active in using the Internet for *commercial* purposes (Liu, Li, & Liu, 2015). According to a study by McKinsey, online retail channels created more new spending in China's underdeveloped regions than in developed ones (Dobbs et al., 2013). If commercial activities are of any guidance, they suggest that the rise of online participation opportunities such as the LLMB will also likely attract more new participants from the less prosperous areas.

If the LLMB can indeed help reduce barriers for lower-class participation, a question that naturally follows is: how would that affect government policies? A large literature on political economy suggests lower-class citizens often hold stronger preferences for redistribution (e.g., Acemoglu & Robinson, 2005; Hibbs, Rivers, & Vasilatos, 1982; Meltzer & Richard, 1981). Thus, our second hypothesis is that *government policies will become more redistributive as the voices of the disadvantaged become better heard*. Although government officials in China are often not held electorally accountable to their citizens, existing studies suggest several reasons why citizen

---

<sup>7</sup>Ministry of Industry of Information Technology, <https://goo.gl/CBTGrS>. Although ownership of personal computer in rural areas is still not as common as in urban areas, the presence of Internet cafes and other public access venues provides inexpensive alternatives.



participation can still lead to policy changes even in the absence of electoral institutions. Participation may, for example, reveal important information about citizen preferences and local conditions (Distelhorst & Hou, 2017; Lorentzen, 2013), which can be used by an official's superior to evaluate his/her performance. Compared to electoral democracies, the Chinese regime may be even more dependent on information conveyed through channels such as the Internet because of the underdevelopment of more conventional democratic institutions. Moreover, responsiveness may also be driven by officials' fear of citizens' collective actions (Chen et al., 2016). To the extent that online grievances are important signals of mass dissatisfaction, authorities may be motivated to make substantive policy concessions in response to these grievances as a way to preempt more disruptive offline actions.

Several pieces of anecdotal evidence suggest that local governments do take online petitions from the LLMB seriously. Shortly after the platform was launched, many local governments formed partnership with the platform and set up specialized agencies to process petitions and regularly report key information from online petitions to high-level decision-makers.<sup>8</sup> When drafting policy documents, such as the Government Work Reports (see below), the central and local authorities have also shown a willingness to seek and incorporate suggestions from online platforms.<sup>9</sup> Moreover, many local politicians use this platform to publicly engage with their constituencies: As of 2015, over 50 provincial party secretaries and governors, the highest-ranking regional leaders in China, have written personal replies to citizen petitions at the LLMB, along with many more city- and county-level leaders.<sup>10</sup>

---

<sup>8</sup>Personal interviews, 2017. See Online Appendix Q for more information about the petition-policy linkage.

<sup>9</sup>The central government, for example, initiated a campaign entitled "I spoke to the Premier" to encourage netizen input to the drafting of the Central Government Work Report in 2014. Many local governments have also made their own initiatives. See "政府工作报告起草组负责人解读报告", *Xinhua News*, <https://goo.gl/d1oDQP>, and "盐田区2016年政府工作报告向网友征求建议", *Shenzhen News*, <https://goo.gl/qwS8SG>.

<sup>10</sup>On average, about 60% of the petitions on LLMB have received replies from local government agencies.

## 4 Analyzing LLMB Petitions: Issues and User Backgrounds

We scraped all the publicly available petitions from the website (~900,000), along with a rich set of information for each petition, such as (user-classified) subject matter, time of posting, the identity of the leader to whom the petition was directed, whether a government reply has been made, and so on. We began with a descriptive analysis of the range of issues raised on the platform as well as petitioners' areas of origins. To do so, we took a random sample of 3,500 petitions and hired two research assistants (RAs) to read through and assign to each petition one of 14 issue labels and 3 location labels based on the content of the petition.<sup>11</sup> Since not all petitions contain sufficient information for us to make unambiguous coding decisions, sample statistics based on coded petitions alone are likely to be biased.<sup>12</sup> To remedy this, we adopted a non-parametric method developed by Hopkins and King (2010) to infer the overall distribution of issues and locations in the entire body of petitions based on the coded sample. The basic intuition of this method is that the frequencies of different words in a corpus can be expressed as the product of (1) the word frequencies in different issue categories and (2) the relative shares of these categories. Using the hand-coded information of word frequencies in different categories (i.e., labels) from the training sample as a proxy for quantity (1), we can back out the label composition that is most likely to generate the aggregate word frequencies observed in the entire body of petitions (i.e., quantity (2)).<sup>13</sup> We provide more details on the estimation procedure in Online Appendix E.

Table 1 displays the estimated proportions for issues and locations, all arranged in descending orders. For issues, the most frequently raised ones include (1) property transaction and management, (2) employment, (3) neighborhood environment, (4) land taking and house demolition, (5) education and health, (6) agriculture production, (7) labor disputes, and (8) social security. With

---

<sup>11</sup>In creating the issue groups, we consulted the output from unsupervised LDA topic models (discussed in Section H, OA). For consistency concerns, the research assistants are instructed to independently go through *all* sampled petitions and we compare their coding results afterwards. The inter-coder reliability is above 80% for both dimensions. When the RAs' coding decisions disagree, we read the petition and determine the proper classification by ourselves.

<sup>12</sup>We could not unambiguously determine the issue topic for about 4% of the petitions and petitioner's residence for about 19% of the petitions. These petitions are coded as missing.

<sup>13</sup>In doing so, we need to make two assumptions. The first is that that individuals who share similar grievances or social backgrounds use similar linguistic patterns, and the second is that the distribution of word for each issue in the training set is representative of the population.

the possible exception of (1),<sup>14</sup> the other seven of the top eight issues all seem to be life problems that are more commonly experienced by the lower class than by the middle or the upper class. Collectively, the seven issues account for about 50% of all petitions.

Table 1: Estimated Distributions for Issues and Petitioners' Backgrounds

| Issue                            | Proportion | Location    | Proportion |
|----------------------------------|------------|-------------|------------|
| Property transaction/management  | 0.211      | Urban       | 0.55       |
| Employment                       | 0.111      | Rural       | 0.32       |
| Neighborhood environment         | 0.099      | Town/suburb | 0.13       |
| Land taking and house demolition | 0.093      |             |            |
| Education/Healthcare             | 0.074      |             |            |
| Agricultural production          | 0.072      |             |            |
| Labor disputes                   | 0.059      |             |            |
| Social security                  | 0.052      |             |            |
| Business                         | 0.046      |             |            |
| Traffic and transportation       | 0.046      |             |            |
| Public safety                    | 0.044      |             |            |
| Infrastructure                   | 0.041      |             |            |
| Hukou                            | 0.026      |             |            |
| Corruption                       | 0.024      |             |            |
| RMSE of 10-fold cross validation | 0.023      |             | 0.057      |

*Note:* This table shows the estimated proportions for petition issues and petitioners' area of residence using semi-automated content analysis. For each model, we report the Root Mean Squared Error, which indicates the average deviation of estimated proportion from the true proportion in a 10-fold cross validation.

The composition of LLMB users can be seen even more clearly as we turn to the results on location labels.<sup>15</sup> While the majority of the petitions still come from urban areas, about 45% of the petitions appear to be filed by users residing in rural and suburban areas, which are usually less developed economically.<sup>16</sup> As a benchmark, it is worth noting that as of 2016, rural residents still

<sup>14</sup>The large proportion of property-related issues is likely to be due to their frequent occurrence in daily life. Our reading suggests that many of the complaints under this category are about malfunctions of facilities in one's apartments/residential compounds.

<sup>15</sup>We determine petitioners' locations of residence based on the content of their petitions, which often mention where they live or even current address. Urban keywords include: residential complex 小区, urban community 社区, residential committee 居委会, urban district 城区, inside a city 城里. Rural: inside a village 村里, village 村庄/乡村, village committee 村委会, in countryside 乡下, mountain village 山村. Suburban: urban village 城中村, rural-urban fringe 城乡结合部, township 镇子, central town 中心镇, city outskirts 城郊, suburb 郊区.

<sup>16</sup>In the context of China, rural residents are in generally poorer than those living in urban areas. In 2008 (the launching year of LLMB), the average disposal income for rural residents was less than 1/3 of that of urban residents' (4761 yuan vs. 15781 yuan). See <https://goo.gl/K3tfSp> (in Chinese).

accounted for only about a quarter of active Internet users in China (CNNIC, 2016). Rural and lower-class users, in other words, are over-represented on the LLMB relative to their share in the Internet-using population.

## **5 The Policy Consequences of Online Participation**

### **5.1 Data on Local Government Policy Priorities**

To analyze the aggregate policy consequences of online participation, we made use of an original panel dataset of government policy priorities created from the text of local Government Work Reports (政府工作报告, GWRs thereafter). GWRs are a form of official communication between the Chinese government and the legislative body at the same level. They are delivered once a year by the head of the government at the annual session of the People's Congress (PC) and have to be formally approved by PC deputies through an anonymous vote. As one of the most important policy documents that local administrations issue every year, GWRs provide comprehensive descriptions of governments' policy blueprints and highlight key socioeconomic targets for the following year. They have to be collectively read and edited by the party standing committees—the core leadership body in local governments—before they are sent out to the legislature for approval.

GWRs have a highly standardized structure: They often begin with a brief description of the overall national and local conditions, followed by a summary of governments' achievements in the past year; the bulk of the document is then devoted to laying out plans and directions for the next year. While the set of major policy areas that needs to be covered in a GWR is usually fixed (e.g., economy, public safety, culture, education, social welfare, etc), the relative amount of emphasis on each area is often subject to political discretion and can reflect local leaders' own policy visions. Newly published GWRs often receive close attention from media and government officials because they contain important information about leadership's policy preferences.

We collected the full text of over 4,400 government reports at both the city and provincial lev-

els between 2000 and 2014,<sup>17</sup> and used a Latent Dirichlet Allocation (LDA) model (Blei, Ng, & Jordan, 2003) to uncover topics (i.e., cluster of words) from the text and estimate their relative proportions in each report. Compared to the traditional, dictionary-based coding methods, a distinct strength of the LDA algorithm is that it clusters words strictly according to their co-occurrence patterns, thus avoiding the arbitrariness and errors in hand coding. This is especially useful for analyzing complex policy documents such as the GWRs, for which commonly agreed upon coding rules do not yet exist. Moreover, since many words can be related to multiple policy areas, human coder may face difficulties in determining to which topic a particular word should be assigned. The LDA model provides a solution to this problem by allowing each word to be associated with multiple topics at different levels of strength.

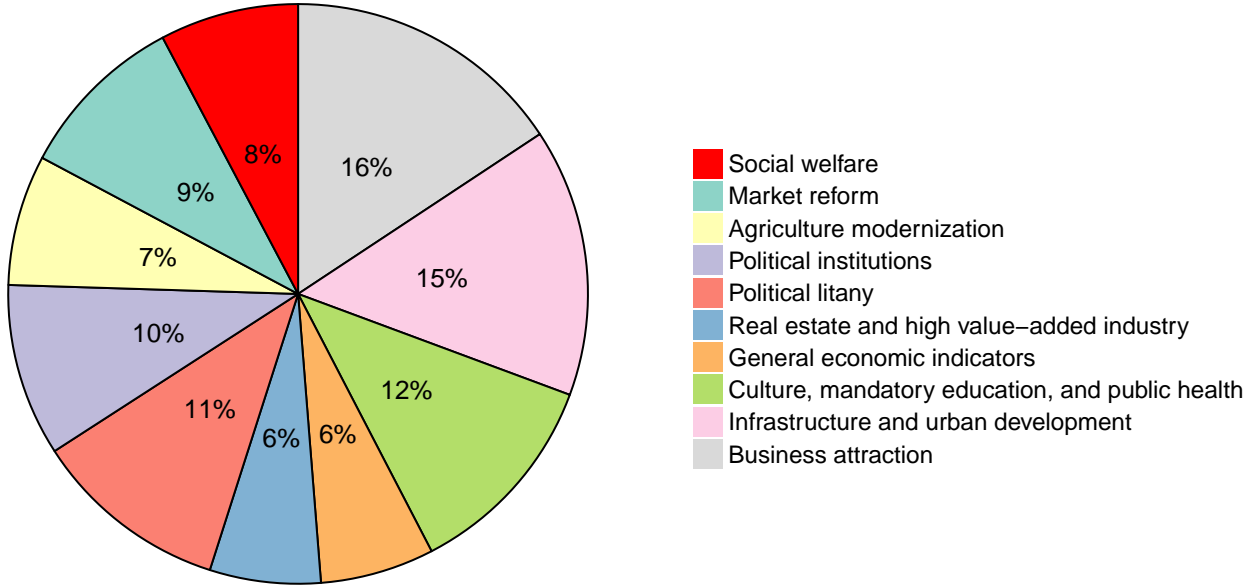
We estimated a 10-topic model as the baseline and a 20-topic model for robustness checks. Figure 1 plots the aggregate distribution of topic shares in the 10-topic model.<sup>18</sup> In light of the previous findings on user background, we choose to focus on a topic for social welfare issues, which is a key concern for lower-class citizens.<sup>19</sup> In this model (as well as in the 20-topic model), we are able to detect a clear and highly coherent topic cluster on social welfare issues, which on average accounts for about 8% of a GWR. High frequency words under this topic include *social safety net* (社会保障), *endowment insurance* (养老保险), *medical insurance* (医疗保险), and *social insurance* (社会保险). According to several post-estimation diagnostics that we performed, the social welfare topic is ranked as one of the highest-quality topics generated by both the 10- and 20-topic LDA models (see Online Appendix F for detailed post-estimation diagnostics). A variance decomposition suggests that within-city variation accounts for about 63% of the total variation in this topic, whereas between-city variation accounts for the rest of the 37%.

<sup>17</sup>Full-text GWR data were collected from government websites and local yearbooks.

<sup>18</sup>The keywords with the highest association with each topic are detailed in Online Appendix F.4. We show that LDA is able to find highly meaningful policy areas. The output from the 20-topic model is available upon request.

<sup>19</sup>In Table A.2, we provide survey-based evidence on the difference in policy preferences for citizens of different socioeconomic backgrounds. Figure A.6 presents results on the effect of LLMB petitions on *all* GWR policy topics.

Figure 1: Aggregated Distribution of GWR Topic Shares, 2000-2014



## 5.2 Empirical Strategy

Our baseline is a fixed-effects model with the following specification:

$$\Delta \text{Welfare Topic Share}_{it+1} = \delta \text{Log Petitions}_{it} + \mathbf{X}_{it}\boldsymbol{\beta} + \phi_i + \tau_t + \epsilon_{it},$$

where  $i$  indexes cities and  $t$  indexes years.<sup>20</sup> The dependent variable,  $\Delta \text{Welfare Topic Share}$ , is the incremental change in the proportion of welfare topic in a city's GWR from the previous year.<sup>21</sup> Because the delivery of the government work reports typically occur at the beginning of a year (January or February), we match each city-year spell with the change in GWR welfare topic in the following year.

The independent variable,  $\text{Log Petitions}_{it}$  is the (logged) aggregate number of petitions about a city in a given year. This includes both petitions filed directly at the city's own message boards and

<sup>20</sup>Our main dataset covers 299 cities for the period of 2008 to 2013. Due to a high level of missingness in GWR data and relatively low usage of the LLMB, we drop three far-flung western provinces (Xinjiang, Tibet, and Qinghai) from the sample. Our results remain robust to including these provinces in the sample (Table A.8).

<sup>21</sup>We choose to use change in welfare topic share as the dependent variable instead of its level to deal with the persistence and non-stationarity of welfare topic share.

those filed at the city's supervising provincial government regarding issues from a specific city.<sup>22</sup> Later, we also use a LDA model to detect distinct topic groups from the entire corpus of petitions and assign each individual petition to a specific topic.<sup>23</sup> This allows us to examine which subset of petitions has the strongest impact on welfare policy change.  $\phi_i$  and  $\tau_t$  represent city and year fixed-effects, respectively. With city fixed-effects, we difference out any variation that is city-specific but time-invariant, and exploit the within-city variation in petition volume and welfare policy. The main regression also includes a set of province-specific linear trends to account for influence from time-variant, province-specific unobserved factors.<sup>24</sup>  $\mathbf{X}_{it}$  is comprised of a rich set of time-variant controls, which are described below.

### Socioeconomic Conditions

The most important potential confounder to our analysis is the intensity of public grievances expressed through other non-Internet channels.<sup>25</sup> Since dissatisfied citizens may concurrently use multiple means to make their grievances heard, changes in government policies may be attributable to their offline actions rather than online participation per se. Prior research has shown that collective protests are one of the most common means for lower-class citizens in China to express their discontent offline. Large-scale protests often receive a good deal of attention from the authority and can sometimes produce major shifts in government policies (Heurlin, 2016).

To address this, we make use of two of the best available datasets. The first one is the Collective Incidents Dataset, compiled by the Institute of Sociology at the Chinese Academy of Social Sciences (CASS). This dataset contains detailed information about major mass protests in China

---

<sup>22</sup>For the second type of petition, we identify their city origins based on posters' IP address and mentioning of city keywords in texts. When a petition mentions multiple city name keywords, we pick the unit that receives most mentioning as a petition's origin city.

<sup>23</sup>This procedure does not take input from the previous aggregate-level analysis. For each petition, the LDA model generates estimated proportions for all 30 topics. We assign each petition to the topic that takes up the highest proportion in that petition. For details, see Online Appendix H.

<sup>24</sup>An alternative specification is to use province-year dummies, which would consume more degrees of freedom in the estimation. Our main result is substantively the same when using this alternative specification.

<sup>25</sup>Although Chinese citizens do not enjoy the right to elect their government officials, prior research has shown that the regime does contain several non-electoral channels for citizens to make their voices heard (e.g., Shi, 1997; Tang, 2016).

between 2007 and 2013, collected from both internal government documents and through extensive searches of domestic and overseas media reports. The second one, *China Strikes* ([www.chinastrikes.crowdmap.com](http://www.chinastrikes.crowdmap.com)), is a crowd-sourced website that focuses specifically on labor-related unrest. For our analysis below, we combine all the unique events from both sources and calculate the frequency of protests for each city-year spell as a control variable.

In addition to protest frequency, we also include controls for a city's population (*Log Population*) and the size of employment (*Log Employment*), as unemployment is a common source of popular grievances and a main driver for welfare expansion. Moreover, because welfare-related spending in China is covered primarily by local government budgets, it is likely to be affected by local economic and fiscal conditions. We thus include *Log GDP*, *GDP Growth Rate*, *Log Fiscal Revenue*, and *Log Fiscal Expenditure* to control for a locality's level and pace of development as well as fiscal capacity.

### **Leadership Characteristics**

The second set of controls are about personal characteristics of local leaders. We include a number of demographic variables (for both the city secretary and the mayor), including age, tenure length, and political connection with the provincial leadership. Age and tenure length, in particular, have been found to be correlated with the career incentives of local leaders as well as their policy preferences. We also include the number of years a city leader has served in a given locality, as prior studies have shown that officials with longer local careers may be more attentive to local interests and spend more on social issues (Persson & Zhuravskaya, 2016).

## **5.3 Baseline Results**

The main results are presented in Table 2. Column 1 presents the most parsimonious model with only two sets of fixed-effects and linear province trends as controls. We see that the (logged) total volume of online petitions is strongly and positively associated with governments' emphases on welfare policies. Specifically, the coefficient indicates that for an average city, a one standard



deviation increase in online petition (~155 more petitions for a median city) is associated with a 0.48 percentage points, or about 17% of a standard deviation, increase in GWR welfare topic share. To put this figure in perspective: in a recent study of policy responsiveness in the United States, Caughey and Warshaw (2017) find that in the non-South, a one standard deviation change in mass liberalism is associated with about 3.7% of a standard deviation immediate change in social policy and 1.4% standard deviation change in economic policy. In another study, M. K. Miller (2015) finds that a one standard deviation loss in electoral vote in electoral autocracies is associated with about 30% of a standard deviation increase education and welfare spending. While such comparison should always be carried out with caution as these studies are based on very different measures of public sentiments, it nonetheless still provides some reassurance that the magnitude of our estimate is within a reasonable range.

Next, we decompose petitions into several more refined topic groups and examine their respective impacts on welfare policy change. We apply a 30-topic LDA model to the petition text and, based on the estimated topics, classify petitions by (1) whether they are from rural or urban areas, and (2) whether they involve issues directly related to one's pocketbook conditions. This gives rise to four distinct groups of petitions: rural petitions about pocketbook issues (RP), urban petitions about pocketbook issues (UP), rural petitions about non-pocketbook issues (RN), and urban petitions about non-pocketbook issues (UN). The details about the classification can be found in Online Appendix H.2. Type RP petitions are those related to rural low-income assistance, land-taking compensations, and financing of basic public goods. Common issues in type UP petitions include medical reimbursement, compensation for house demolition, wage disputes with employers, and unemployment benefits. Non-pocketbook issues include environmental degradation for rural areas and traffic control, public safety, education access, and property management for urban areas. Our expectation is that changes in welfare topics should be more strongly associated with types RP and UP petitions than with the other two types.

Table 2: Main Results

|                                      | $\Delta$ Welfare Topic at $t + 1$ (10-topic) |                       |                       |                       |                       |
|--------------------------------------|--|-----------------------|-----------------------|-----------------------|-----------------------|
|                                      | (1)  | (2)                   | (3)                   | (4)                   | (5)                   |
| Log petitions                        | 0.0048**<br>(0.0021)                         |                       |                       |                       |                       |
| Log petitions (pocketbook)           |  | 0.0085***<br>(0.0029) |                       |                       |                       |
| Log petitions (non-pocketbook)       |  | -0.0025<br>(0.0028)   |                       |                       |                       |
| Log petitions (rural pocketbook)     |  |                       | 0.0073***<br>(0.0021) | 0.0075***<br>(0.0022) | 0.0076***<br>(0.0022) |
| Log petitions (urban pocketbook)     |  |                       | 0.0019<br>(0.0022)    | 0.0022<br>(0.0023)    | 0.0021<br>(0.0023)    |
| Log petitions (rural non-pocketbook) |  |                       | -0.0005<br>(0.0016)   | -0.0004<br>(0.0016)   | -0.0005<br>(0.0016)   |
| Log petitions (urban non-pocketbook) |  |                       | -0.0023<br>(0.0028)   | -0.0026<br>(0.0028)   | -0.0026<br>(0.0029)   |
| City and year fixed-effects          | ✓  | ✓                     | ✓                     | ✓                     | ✓                     |
| Province-specific trends             | ✓  | ✓                     | ✓                     | ✓                     | ✓                     |
| Economic controls                    |  |                       |                       | ✓                     | ✓                     |
| Leadership controls                  |  |                       |                       |                       | ✓                     |
| R <sup>2</sup>                       | 0.02   | 0.03                  | 0.03                  | 0.04                  | 0.04                  |
| Number of Cities                     | 299  | 299                   | 299                   | 298                   | 297                   |
| Observations                         | 1656   | 1656                  | 1656                  | 1625                  | 1624                  |

**Note:** This table shows the results from OLS regressions. The dependent variables are incremental increase in the share of social welfare topic in government work reports. Socioeconomic controls include *Log Protest*, *Log Employment*, *Log Population*, *GDP Growth*, *Log Fiscal Revenue*, and *Log Fiscal Expenditure*. Leadership controls include (for both party secretary and mayor): *Age*, *Tenure*, and *Years of Local Service*, as well as an indicator for any city leader connected to the incumbent provincial secretary. Robust standard errors clustered at city level are reported in parentheses.

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$  (two-tailed test)

Columns 2 through 5 of Table 2 report the results from using the more refined petition groupings. Column 2 only makes the distinction between petitions concerning pocketbook and non-pocketbook issues, and Column 3 further distinguishes between petitions from rural and urban areas. Columns 4 and 5 incrementally add to the specification of Column 3 socioeconomic and leadership controls. The results largely confirm our expectation: Petitions concerning pocketbook issues appear to matter more for welfare policy change than those with non-pocketbook demands, and pocketbook petitions from rural areas appear to have the strongest association with changes in welfare topic share among the four types of petitions. For pocketbook petitions from urban areas, the coefficient is still positive but much smaller and less precisely estimated. One explanation

for the weaker association may be that active participation by the middle class in urban areas has diverted some government attention from the lower class (as evidenced by the high frequency of property-related complaints).

## **5.4 Robustness and Endogeneity**

We conduct several additional robustness checks on the main results. In the interest of space, we only provide a brief summary here and leave the details to the Online Appendix. We find that our main results are unchanged when we use welfare topic proportion generated by an alternative, 20-topic LDA model as the dependent variable (Table A.7). Our results are also robust to including or excluding cities with special political status or distinct ethnic compositions (Tables A.8). Moreover, we show that our results are not sensitive to removing cities with active local online petition platforms (Table A.10).

We also take special care to address the issue of endogeneity. One major concern, for example, is that the volume of online petition may itself be a result of prior government welfare policies. To address this concern, we conduct several Granger-style tests. We find little evidence to support the reverse link: GWR welfare topic shares are not strongly correlated with petition volumes in subsequent years (Online Appendix N). Another concern is that both policies and online participation may be driven by some unobserved time-varying events, such as a shift in societal preference for social welfare. We address this issue through an instrumental variables (IV) analysis, which we detail in Online Appendix O. The IV results are largely consistent with the OLS ones.

## **5.5 Effects of Online Participation on Substantive Outcomes**

The preceding analyses have demonstrated that the volume of LLMB petitions, especially those about rural, pocketbook issues, has a measurable impact on the emphasis of social welfare policies in government work reports. However, one may still question whether changes in policy rhetoric reflect actual changes in governing priorities. To address this issue, we examine the effect of online participation on more substantive policy outcomes. Our specific focus here is the Minimum

Living Standard Guarantee Scheme (*dibao*), which is a major cash-based social assistance program targeted at the poor. Although this program is not the only welfare program that the government implements, we choose to focus on it because it has the best available data and is unambiguously an issue that concerns the lower class. Researchers have shown that local governments typically have a lot of discretion in designing and implementing their own *dibao* programs (Solinger & Hu, 2012).<sup>26</sup> We thus expect that they will expand this program when pocketbook demands from lower-class groups become more salient.

We collect city-level data on the coverage of the rural *dibao* program from the website of the Ministry of Civil Affairs, and correlate them with type RP petitions (along with other types). Table 3 displays the regression results for three metrics: individual coverage, family coverage, and total spending. We can see that RP petitions have a positive and significant impact on the implementation of rural *dibao* across all these metrics. The instrumental variables analysis yields largely similar, if not stronger, results (Table A.14). These results seem to suggest that online participation does have substantive redistributive consequences that go beyond policy rhetoric.

---

<sup>26</sup>The *dibao* program is jointly funded by central and local governments, and local governments play an especially prominent role in raising money for the rural *dibao*, which is the focus of our empirical analysis here. Other responsibility of local governments include setting the maximum eligible income for local *dibao* applicants, approving *dibao* applications, disbursing the assistance, and monitoring the use of the fund (e.g., Bai & Gu, 2018).

Table 3: Change in Substantive Outcomes: Rural Minimum Living Standard Guarantee Scheme (*dibao*)

|                                      | Coverage (individual) | Coverage (family)     | Total spending        |
|--------------------------------------|-----------------------|-----------------------|-----------------------|
|                                      | (1)                   | (2)                   | (3)                   |
| Log petitions (rural pocketbook)     | 0.0167**<br>(0.0078)  | 0.0371***<br>(0.0127) | 0.0287***<br>(0.0102) |
| Log petitions (urban pocketbook)     | -0.0002<br>(0.0129)   | 0.0008<br>(0.0150)    | -0.0170<br>(0.0155)   |
| Log petitions (rural non-pocketbook) | -0.0101<br>(0.0062)   | -0.0095<br>(0.0080)   | -0.0088<br>(0.0095)   |
| Log petitions (urban non-pocketbook) | -0.0108<br>(0.0125)   | -0.0202<br>(0.0167)   | -0.0106<br>(0.0146)   |
| City and year fixed-effects          | ✓                     | ✓                     | ✓                     |
| Number of Cities                     | 301                   | 301                   | 301                   |
| Observations                         | 1762                  | 1762                  | 1763                  |

**Note:** The table reports results using several implementation outcomes of the rural *dibao* program as the dependent variable. The specification is based on Column 5 of Table 2.

Robust standard errors clustered at city level are reported in parentheses.

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$  (two-tailed test)

## 6 Concluding Remarks

As one of the most influential technological inventions of the 20th century, the Internet has been seen by many as having the potential to give citizens a louder voice in the political process. However, concrete evidence on whether and how Internet-based participation can affect government policies remains limited to date. This study addresses this gap by presenting new evidence on the policy consequences of online participation from China's largest electronic petition platform. Contrary to the prevailing view that in developing countries the Internet is still a luxury technology enjoyed mainly by the middle- and upper-class elites (P. Norris, 2001), our results suggest that when broadband infrastructure is made nearly universal, those from lower-class backgrounds will actively take advantage of new online channels to articulate their interests to the authorities, and that their participation can lead local governments to place higher priority on social welfare issues in both rhetoric and actual policy implementation.

Our study speaks to the broader literature on the relationship between the Internet and governance. Existing research has found that the Internet can help strengthen government account-

ability (Besley & Burgess, 2002) and reduce corruption (Bailard, 2009), but may also suppress voter turnout (Campante, Durante, & Sobbrío, 2013) and increase ideological polarization (Lelkes, Sood, & Iyengar, 2015, 1). On the issue of political inequality, more specifically, empirical findings from survey-based studies in advanced democracies (the U.S in particular) tend to support the view that the Internet will reinforce, rather than alleviate, the participation and influence gaps in the offline world (P. Norris, 2001; Schlozman, Verba, & Brady, 2010). Our findings suggest that the conclusion drawn from the experience of developed countries may not necessarily apply to a developing context where the opportunities for conventional forms of participation are more biased and limited. We provide evidence that supports the Internet's potential to mitigate participation inequality among citizens and show that bringing in new voices into politics can substantively change governments' policy priorities. While we certainly need to be cautious to not over-generalize findings from a particular platform, it is worth noting that some of the key features that made the platform effective, including operational independence, ease of access, and transparency, are not rare qualities in the cyberspace. As many other countries and regions have started to adopt similar electronic petition platforms, a promising direction of future research will be to investigate whether a similar equalizing impact also exists in other country settings.

This study also contributes to a nascent but rapidly growing literature on non-electoral responsiveness. Several recent studies have shown that in some non-Western political systems, and China in particular, inquiries and demands made by individual citizens receive replies from governments at a comparable rate as in electoral democracies (Chen et al., 2016; Distelhorst & Hou, 2017), and that individual politicians are generally attentive to citizen opinions (Meng, Pan, & Yang, 2017; Truex, 2016). So far, however, there is still limited evidence on whether, in the absence of formal electoral accountability, the preferences and demands expressed by citizens will be systematically taken into account in government *policy making*—a more fundamental criterion of political responsiveness according to classical writings on this concept (Dahl, 1971; W. E. Miller & Stokes, 1963). Our findings suggest the existence of a form of responsiveness in China that matches with its canonical definition in the democratic context.

Finally, our analysis advances a more nuanced view on the relationship between the Internet and the durability of the Chinese regime. While previous studies tend to portray a largely confrontational picture, focusing on either the subversive capacity of the Internet or the regime's efforts to control and manipulate the cyberspace (King, Pan, & Roberts, 2013, 2017), our findings suggest that the interaction between the two is not always zero-sum: the democratic potential of the Internet may be harnessed by the government to fulfill important governing functions. As such, the development of Internet may contribute to the regime's vitality by allowing it to improve the quality of governance without making more radical changes to its political institutions.

## References

- Acemoglu, D., & Robinson, J. A. (2005). *Economic origins of dictatorship and democracy*. Cambridge University Press.
- Alesina, A., & Ferrara, E. L. (2005). Preferences for redistribution in the land of opportunities. *Journal of Public Economics*, 89(5–6), 897–931.
- Bai, C., & Gu, X. (2018). Horizontal inequality in china's social safety net. *Chinese Public Administration*, 1, 109–115. in Chinese.
- Bailard, C. S. (2009). Mobile phone diffusion and corruption in africa. *Political Communication*, 26(3), 333–353.
- Besley, T., & Burgess, R. (2002). The political economy of government responsiveness: Theory and evidence from india. *The Quarterly Journal of Economics*, 117(4), 1415–1451.
- Bimber, B. (1998). The internet and political transformation: Populism, community, and accelerated pluralism. *Polity*, 31(1), 133–160.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3, 993–1022.
- Campante, F. R., Durante, R., & Sobbrío, F. (2013). *Politics 2.0: The multifaceted effect of broadband internet on political participation* (Working Paper No. 19029). National Bureau of Economic Research. Retrieved from <https://goo.gl/6vShfB>
- Castells, M. (1996). *The rise of the network society: The information age: Economy, society and culture*. Wiley.
- Caughey, D., & Warshaw, C. (2017). Policy preferences and policy change: Dynamic responsiveness in the american states, 1936–2014. *American Political Science Review*, 112(02), 249–266.
- Chadwick, A., & May, C. (2003). Interaction between states and citizens in the age of the internet: "e-government" in the united states, britain, and the european union. *Governance*, 16(2), 271–300.
- Chen, J., Pan, J., & Xu, Y. (2016). Sources of authoritarian responsiveness: A field experiment in china. *American Journal of Political Science*, 60(2), 383–400.
- CNNIC. (2016). 2015 report on rural internet development. Retrieved from <https://goo.gl/aWa8Zr>
- Dahl, R. A. (1971). *Polyarchy: Participation and opposition*. New Haven, CT: Yale University Press.
- Distelhorst, G., & Hou, Y. (2017). Constituency service under nondemocratic rule: Evidence from china. (Vol. 79, 3, pp. 1024–1040).
- Dobbs, R., Chen, Y., Orr, G., Manyika, J., Chui, M., & Chang, E. (2013). *China's e-tail revolution*. McKinsey Global Institute. Retrieved from <https://goo.gl/zQKdao>
- Dutton, W. H. (2009). The fifth estate emerging through the network of networks. *Prometheus*, 27(1), 1–15.
- Goodman, D. S. G. (2014). *Class in contemporary china*. Polity Press.
- Guo, G. (2007). Organizational involvement and political participation in china. *Comparative Political Studies*, 40(4), 457–482.
- Heurlin, C. (2016). *Responsive authoritarianism in china: Land, protests, and policy making*. Cambridge University Press.



- Hibbs, J., Douglas A., Rivers, R. D., & Vasilatos, N. (1982). On the demand for economic outcomes: Macroeconomic performance and mass political support in the united states, great britain and germany. *Journal of Politics*, 44, 426–462.
- Hopkins, D., & King, G. (2010). A method of automated nonparametric content analysis for social science. *American Journal of Political Science*, 54(1), 229–247.
- Kim, S., & Lee, J. (2012). E-participation, transparency, and trust in local government. *Public Administration Review*, 72(6), 819–828.
- King, G., Pan, J., & Roberts, M. E. (2013). How censorship in china allows government criticism but silences collective expression. *American Political Science Review*, 107(2), 326–343.
- King, G., Pan, J., & Roberts, M. E. (2017). How the chinese government fabricates social media posts for strategic distraction, not engaged argument. *American Political Science Review*, 111(3), 484–501.
- Lelkes, Y., Sood, G., & Iyengar, S. (2015). The hostile audience: The effect of access to broadband internet on partisan affect. *American Journal of Political Science*, 61, 5–20.
- Linde, J., & Karlsson, M. (2013). The dictator's new clothes: The relationship between e-participation and quality of government in non-democratic regimes. *International Journal of Public Administration*, 36(4), 269–281.
- Lindner, R., & Riehm, U. (2011). Broadening participation through e-petitions? an empirical study of petitions to the german parliament. *Policy & Internet*, 3(1), 63–85.
- Liu, C., Li, J., & Liu, J. (2015). Rural e-commerce and new model of rural development in china: A comparative study of "taobao village" in jiangsu province. *Asian Agricultural Research*, 7(11), 35–46.
- Lorentzen, P. (2013). Regularizing rioting: Permitting public protest in an authoritarian regime. *Quarterly Journal of Political Science*, 8(3), 127–158.
- Marien, S., Hooghe, M., & Quintelier, E. (2010). Inequalities in non-institutionalised forms of political participation: A multi-level analysis of 25 countries. *Political Studies*, 58(1), 187–213.
- Meltzer, A., & Richard, S. F. (1981). A rational theory of the size of government. *Journal of Political Economy*, 89(5), 914–927.
- Meng, T., Pan, J., & Yang, P. (2017). Conditional receptivity to citizen participation. 50(4), 399–433.
- Miller, M. K. (2015). Elections, information, and policy responsiveness in autocratic regimes. *Comparative Political Studies*, 48(6), 691–727.
- Miller, W. E., & Stokes, D. E. (1963). Constituency influence in congress. *The American Political Science Review*, 57(1), 45–56.
- Norris, D. F., & Reddick, C. G. (2012). Local e-government in the united states: Transformation or incremental change? *Public Administration Review*, 73(1), 165–175.
- Norris, P. (2001). *Digital divide: Civic engagement, information poverty, and the internet world-wide*. Cambridge University Press.
- Oreglia, E. (2015). The “sent-down” internet: Using information and communication technologies in rural china. *Chinese Journal of Communication*, 8(1), 1–6.
- Pan, J. (2017). How chinese officials use the internet to construct their public image. *Political Science Research and Methods*, 1–17.
- Persson, P., & Zhuravskaya, E. (2016). The limits of career concerns in federalism: Evidence from china. *Journal of the European Economic Association*, 14(2), 338–374.

- Royo, S., Yetano, A., & Acerete, B. (2013). E-participation and environmental protection: Are local governments really committed? *Public Administration Review*, 74(1), 87–98.
- Schlozman, K. L., Verba, S., & Brady, H. E. (2010). Weapon of the strong? participatory inequality and the internet. *Perspectives on Politics*, 8(2), 487–509.
- Shi, T. (1997). *Political participation in beijing*. Cambridge MA: Harvard University Press.
- Solinger, D. J., & Hu, Y. (2012). Welfare, wealth and poverty in urban china: The dibao and its differential disbursement. *The China Quarterly*, 211, 741–764.
- Stromseth, J. R., Malesky, E. J., & Gueorguiev, D. D. (2017). *China's governance puzzle: Enabling transparency and participation in a single-party state*. Cambridge University Press.
- Tang, W. (2016). *Populist authoritarianism: Chinese political culture and regime sustainability*. Oxford University Press.
- Torres, L., Pina, V., & Acerete, B. (2006). E-governance developments in european union cities: Reshaping government's relationship with citizens. *Governance*, 19(2), 277–302.
- Truex, R. (2016). *Making autocracy work*. Cambridge University Press.
- Tsai, L. L., & Xu, Y. (Forthcoming). Outspoken insiders: Political connections and citizen participation in authoritarian china. *Political Behavior*.
- United Nations. (2016). *E-government survey 2016: E-government in support of sustainable development*. Department of Economic and Social Affairs. Retrieved from <https://goo.gl/bsKs9h>
- Verba, S., & Nie, N. H. (1972). *Participation in america: Political democracy and social equality*. Evanston, London.
- Ward, S., Gibson, R., & Lusoli, W. (2003). Online participation and mobilisation in britain: Hype, hope and reality. *Parliamentary Affairs*, 56(4), 652–668.
- Wong, W., & Welch, E. (2004). Does e-government promote accountability? a comparative analysis of website openness and government accountability. *Governance*, 17(2), 275–297.
- Yang, G. (2009). *The power of the internet in china: Citizen activism online*. Columbia University Press.

# Online Appendix (Not for Publication)

|          |  |             |
|----------|--|-------------|
| <b>A</b> | <b>Electronic Government Petition Platforms around the World</b>                           | <b>A-3</b>  |
| <b>B</b> | <b>Survey Evidence on Policy Preferences across Socioeconomic Status</b>                   | <b>A-3</b>  |
| <b>C</b> | <b>A Snap Shot of LLMB's Interface</b>   | <b>A-6</b>  |
| <b>D</b> | <b>Comparing Petition Activities on LLMB vs. Local Platforms</b>                           | <b>A-7</b>  |
| <b>E</b> | <b>Details on Non-Parametric Content Analysis of Petition Issues and Users Backgrounds</b> | <b>A-8</b>  |
|          | E.1 Analysis Procedure . . . . .   | A-8         |
| <b>F</b> | <b>Details on Estimating Government Policy Priorities</b>                                  | <b>A-9</b>  |
|          | F.1 Methodology . . . . .  | A-9         |
|          | F.2 Estimation Procedure . . . . .   | A-10        |
|          | F.3 Post-Estimation Diagnostics . . . . .  | A-12        |
|          | F.4 Estimated Government Work Report Topics . . . . .                                      | A-15        |
|          | F.5 A Sample of GWR with Welfare Topics Highlighted . . . . .                              | A-18        |
| <b>G</b> | <b>Geographic Distribution of Online Petitions</b>   | <b>A-19</b> |
| <b>H</b> | <b>Details on Estimating and Classifying Online Petitions using LDA Models</b>             | <b>A-20</b> |
|          | H.1 Estimating Petition Topics . . . . .   | A-20        |
|          | H.2 Details about Topic Classification . . . . .   | A-22        |
| <b>I</b> | <b>Data on Collective Protests</b>   | <b>A-26</b> |
| <b>J</b> | <b>Validating the Link between Petition and Rural Participants</b>                         | <b>A-28</b> |
| <b>K</b> | <b>Baseline Results with All Coefficients</b>  | <b>A-30</b> |
| <b>L</b> | <b>Effect of Petition on All GWR Policy Topics</b>   | <b>A-31</b> |
| <b>M</b> | <b>Robustness Checks</b>   | <b>A-32</b> |
|          | M.1 Baseline Estimates without City Fixed-Effects . . . . .                                | A-32        |
|          | M.2 Alternative Dependent Variable: GWR Welfare Topic from a 20-Topic Model . . .          | A-33        |
|          | M.3 Estimation on Different Samples . . . . .  | A-33        |
|          | M.4 Addressing the Presence of Local Petition Platforms . . . . .                          | A-35        |
| <b>N</b> | <b>Granger-Style Tests on Reverse Causality</b>  | <b>A-37</b> |
| <b>O</b> | <b>Instrumental Variables Analysis</b>   | <b>A-39</b> |
|          | O.1 Analysis Procedure . . . . .   | A-39        |
|          | O.2 Evidence on Exclusion Restriction . . . . .  | A-42        |

|          |  |             |
|----------|--|-------------|
| O.3      | Robustness Checks and Placebo Tests . . . . .              | A-44        |
| <b>P</b> | <b>Extensions</b>  | <b>A-47</b> |
| P.1      | Effects of Online Petitions on Spending Patterns . . . . . | A-47        |
| P.2      | Effects of Participation over Time . . . . .               | A-48        |
| P.3      | Geographic Variation in Responsiveness . . . . .           | A-49        |
| <b>Q</b> | <b>Fieldwork Evidence on the Petition-Policy Link</b>      | <b>A-50</b> |

## A Electronic Government Petition Platforms around the World

Table A.1: A List of Electronic Petition Platforms Created by National/Regional Governments

| Country/region | Creator<br>Organization(special<br>name, if any)                               | Year of es-<br>tablishment | Website   |
|----------------|--|----------------------------|---|
| United Kingdom | Scottish Parliament  | 1999                       | <a href="http://www.parliament.scot/gettinginvolved/petitions">http://www.parliament.scot/gettinginvolved/petitions</a>                 |
| Australia      | Queensland Parliament  | 2002                       | <a href="https://www.parliament.qld.gov.au/work-of-assembly/petitions">https://www.parliament.qld.gov.au/work-of-assembly/petitions</a> |
| South Korea    | Korean National Govern-<br>ment ( <i>E-People</i> )                            | 2005                       | <a href="http://www.epeople.go.kr/">http://www.epeople.go.kr/</a>   |
| United Kingdom | Prime Minister's Of-<br>fice ( <i>Downing Street</i><br><i>E-Petitions</i> )   | 2006-2010                  | <a href="http://petitions.number10.gov.uk">http://petitions.number10.gov.uk</a>   |
| Germany        | German Bundestag   | 2007                       | <a href="https://epetitionen.bundestag.de/epet/peteinreichen.html">https://epetitionen.bundestag.de/epet/peteinreichen.html</a>         |
| China          | Chinese Communist Party<br>( <i>Local Leader Message</i><br><i>Boards</i> )    | 2008                       | <a href="http://liuyan.people.com.cn">http://liuyan.people.com.cn</a>   |
| United Kingdom | UK Parliament  | 2011                       | <a href="https://petition.parliament.uk">https://petition.parliament.uk</a>   |
| United States  | White House ( <i>We the Peo-<br/>ple</i> )                                     | 2011                       | <a href="https://petitions.whitehouse.gov">https://petitions.whitehouse.gov</a>   |
| Russia         | Russian Federal Gov-<br>ernment ( <i>Russian Public</i><br><i>Initiative</i> ) | 2013                       | <a href="https://www.roi.ru">https://www.roi.ru</a>   |
| Ukraine        | Cabinet of Ministers   | 2016                       | <a href="https://petition.kmu.gov.ua">https://petition.kmu.gov.ua</a>   |

## B Survey Evidence on Policy Preferences across Socioeconomic Status

This section draws on data from the 2008 and 2009 China Citizenship Attitude Surveys (CCAS) to provide evidence on how Chinese citizens' policy preferences vary by their socioeconomic status. Specifically, we utilize a questionnaire item (D10), which asks the respondents to pick an issue that they think is the "most important" from the following 4 options:

1. Maintaining domestic order
2. Give people more voice in government decision-making

3. Protect civil liberty

4. Improve welfare of the poor

We estimate a multinomial logit model where individual choices are regressed on log family income along with other covariates. Table A.2 presents the results. The results suggest that both family income and rural residency are highly significant predictors of preferred policy priorities. Compared to the urban, wealthy groups, rural and lower-income respondents are much more likely to regard welfare improvement for the poor as the most important issue. Consistent with the pattern, we also find that the probability of choosing “don’t know” decreases as income rises. These results thus support our claim that the rich and the poor possess distinct preferences for policy issues in contemporary China.

Table A.2: Different Demands from High- and Low-Income Individuals

| Baseline: public order          | DV: Most Important Task for Government |                   |                      |                       |
|---------------------------------|--|-------------------|----------------------|-----------------------|
|                                 | Voice on policy                        | Freedom           | Improve welfare      | Don't know            |
| Log family income               | 0.147<br>(0.092)                       | -0.042<br>(0.068) | -0.140***<br>(0.045) | -0.349***<br>(0.078)  |
| Rural residency                 | -0.004<br>(0.202)                      | 0.104<br>(0.168)  | 0.343***<br>(0.116)  | 1.190***<br>(0.304)   |
| Age                             | -0.008<br>(0.007)                      | 0.001<br>(0.005)  | 0.006*<br>(0.003)    | 0.035***<br>(0.006)   |
| Female                          | 0.143<br>(0.152)                       | 0.168<br>(0.118)  | 0.455***<br>(0.075)  | 1.285***<br>(0.157)   |
| College education               | -0.247<br>(0.447)                      | -0.170<br>(0.342) | -0.120<br>(0.241)    | -14.079***<br>(0.333) |
| CCP membership (self or family) | -0.103<br>(0.190)                      | -0.119<br>(0.149) | -0.268***<br>(0.096) | -0.793***<br>(0.226)  |
| Employed in state sector        | -0.089<br>(0.259)                      | -0.108<br>(0.197) | -0.109<br>(0.138)    | -0.697*<br>(0.411)    |
| Constant                        | -2.607**<br>(1.073)                    | -0.708<br>(0.788) | 0.864*<br>(0.511)    | -1.618*<br>(0.980)    |
| Province dummy                  | ✓                                      | ✓                 | ✓                    | ✓                     |
| Year dummy                      | ✓                                      | ✓                 | ✓                    | ✓                     |
| Observations                    | 4133                                   | 4133              | 4133                 | 4133                  |

**Note:** This table shows results from a multinomial logit model where the dependent variable is the respondent's choice of the "most important task for government to accomplish". The result suggests that family income is negatively associated with the tendency to select welfare improvement as the most important task for the government. In other words, poor individuals care much more about social welfare issues than wealthy ones.

Data: China Citizenship Attitude Survey (2008 & 2009)

Robust standard errors clustered at individual level are reported in parentheses.

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$  (two-tailed test)

## C A Snap Shot of LLMB's Interface

Figure A.1: A Snapshot of the LLMB

(a) Entry Interface

**Provincial Leader Message Boards**

|  |   |
|--|---|
| <p>Provincial secretary of Henan: Xie Fuzhan<br/>河南省委书记谢伏瞻 <a href="#">查看简历</a></p> <p>年度总留言量: 10061 条 年度公开回复量: 7984 条<br/>Total annual message: 10061 Total reply: 7984</p> <p><a href="#">我要留言</a><br/>“I want to leave a message”</p> | <p>Governor of Henan: Chen Run'er<br/>河南省省长陈润儿 <a href="#">查看简历</a></p> <p>年度总留言量: 3221 条 年度公开回复量: 2557 条</p> <p><a href="#">我要留言</a><br/>“I want to leave a message”</p> |
|--|---|

**City Leader Message Boards**

|   |   |   |  |   |
|---|---|---|--|---|
| <p>郑州市<br/>年度总留言量: 6481 条<br/>年度公开回复量: 5809 条</p> <p><b>Zhengzhou City</b><br/>Total annual message: 6481<br/>Total reply: 5809</p> | <p>开封市<br/>年度总留言量: 2075 条<br/>年度公开回复量: 1094 条</p> | <p>洛阳市<br/>年度总留言量: 2009 条<br/>年度公开回复量: 1846 条</p> | <p>平顶山市<br/>年度总留言量: 5523 条<br/>年度公开回复量: 4259 条</p> | <p>安阳市<br/>年度总留言量: 1485 条<br/>年度公开回复量: 1060 条</p> |
| <p>鹤壁市<br/>年度总留言量: 224 条<br/>年度公开回复量: 49 条</p>  | <p>新乡市<br/>年度总留言量: 3321 条<br/>年度公开回复量: 2784 条</p> | <p>焦作市<br/>年度总留言量: 552 条<br/>年度公开回复量: 330 条</p>   | <p>濮阳市<br/>年度总留言量: 5735 条<br/>年度公开回复量: 4802 条</p>  | <p>许昌市<br/>年度总留言量: 481 条<br/>年度公开回复量: 0 条</p>     |

(b) Message-viewing Interface

您的留言位置: 地方领导留言板 > 河南省 > 河南省委书记谢伏瞻

请输入标题, 不超过22字 Writing the title (no more than 22 characters) [主题领域](#)

主题类别

请输入留言内容 (字数不得少于20字, 不得超过1000字)

Area for writing the message

添加图片  文件尺寸: 小于 500 kb  
可用扩展名: jpg, bmp, gif, png, jpeg

[上传附件](#) Upload attachments [我已阅读并同意《地方领导留言板》管理条约](#)

联系方式  真实姓名  联系电话  仅供工作人员查看, 不对外公开

[提交留言](#)

“Submit the message”

[全部](#) [未回复](#) [办理中](#) [已回复](#)

**为民做主** | 未回复

匿名网友 2017-05-15 09:31

*Uphold Justice for the People* (from anonymous user)  
Dear Respected Leader: Hello! Sorry for bothering you. I want to reflect a problem that I encountered. My name is Wang Weidong and I live in Songquan Village. The village committee forcibly sold 1.6 mu of my farm land without informing my family. I have no other recourse but to ask you to help us.

尊敬的领导, 你们好! 在你百忙之中, 给你添麻烦实感无奈, 向你反映一个问题。我叫王卫东, 住今是街道宋园村委, 只因在我们全家都不知情的情况下, 我1.6亩责任田被村委会强行卖掉, 去村委反映, 村委干部却说让我去告状, 我现在没有办法, 才来向领导求助, 请领导帮帮我们。 [\[查看全文\]](#)

---

**三门峡陕州区购房补贴啥时候能落实到位**

匿名网友 2017-05-15 09:23

*When can I Get My House-Purchasing Subsidy?* (from anonymous user)  
I submitted application for house-purchasing subsidy on Feb 20 2017. So far I haven't heard any progress. I call the relevant departments and got nothing. I just wanted to know when I will be able to receive the subsidy!!!! [\[Click to see the full text\]](#)

本人于2017年2月20号左右提交了陕州区购房补贴的资料, 到目前为止没有任何进展, 打电话咨询相关部门, 一问三不知, 本人就想知道, 现在进展如何, 到哪一步了, 啥时候能补贴到位!!!! [\[查看全文\]](#)



## D Comparing Petition Activities on LLMB vs. Local Platforms

It is important to acknowledge that while the LLMB is the largest online petition platform in China during the period we analyze, it is not the only platform available. Aside from LLMB, many local governments have created their websites to receive citizen complaints. This section provides some evidence on the relationship between petitions on LLMB and local platforms. We scrap petitions from three local websites where the data are publicly available: Sichuan (<https://ly.scol.com.cn>), Changsha (<http://wlwz.changsha.gov.cn/webapp/cs/email/index.jsp>) and Nanjing (<http://www.njbbs.gov.cn/>). Sichuan is a provincial level unit located in southwest China whereas Changsha and Hunan are two prefecture-level units, located in central and eastern parts of China, respectively. We calculate the total number of online petitions filed on these websites each day and regress it on daily petition volumes from the corresponding LLMB platforms.<sup>27</sup> The results are displayed in Table A.3. We can see in all three localities, there exist strong and contemporaneous correlations between LLMB and local petitions. On average, one petition on LLMB is associated with about 0.4 to 1.8 petitions on local sites. These patterns provide suggestive evidence that petition activities that we observe on LLMB are broadly representative of online petition activities on local websites.

---

<sup>27</sup>The observations are somewhat different for the three localities because the local platforms are established on different dates.

Table A.3: Correlation in Petition Volumes: LLMB vs. Local Platforms

|                                     | DV: Daily Petition in Local Platforms |                       |                       |                       |
|-------------------------------------|---------------------------------------|-----------------------|-----------------------|-----------------------|
|                                     | (1)<br>All                            | (2)<br>Sichuan        | (3)<br>Changsha       | (4)<br>Nanjing        |
| Daily petitions at LLMB ( $t$ )     | 0.3826***<br>(0.0223)                 | 0.3836***<br>(0.0222) | 0.8564***<br>(0.2399) | 1.8410***<br>(0.5969) |
| Daily petitions at LLMB ( $t - 1$ ) | -0.0083<br>(0.0243)                   | -0.0114<br>(0.0243)   | 0.7981***<br>(0.2480) | 1.6483***<br>(0.6061) |
| Daily petitions at LLMB ( $t - 2$ ) | -0.0462**<br>(0.0235)                 | -0.0444*<br>(0.0235)  | 0.3729<br>(0.2294)    | 0.1216<br>(0.6075)    |
| Daily petitions at LLMB ( $t - 3$ ) | -0.0486**<br>(0.0196)                 | -0.0447**<br>(0.0195) | 0.3297<br>(0.2127)    | -0.9151<br>(0.5784)   |
| R <sup>2</sup>                      | 0.36                                  | 0.51                  | 0.02                  | 0.10                  |
| Observations                        | 9537                                  | 3211                  | 3365                  | 2961                  |

**Note:** This table shows the correlation in daily petition volumes between LLMB and several local petition platforms (Sichuan, Changsha, and Nanjing). The estimates show that petition activities on local platforms track closely with petitions on LLMB. Robust standard errors are reported in parentheses.

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$  (two-tailed test)

## E Details on Non-Parametric Content Analysis of Petition Issues and Users Backgrounds

### E.1 Analysis Procedure

To create the training dataset, we hire two research assistants (RAs) to read through a random sample of 3,500 petitions and code each petition’s issue and location. In creating the issue labels, we consulted the output from unsupervised LDA topic models (discussed in Section H, OA) and made additional adjustments based on a close reading of hundreds of petitions. A copy of the coding manual is available at <https://www.dropbox.com/s/ufcd4imrwqbo2my/Coding%20Manual.docx?dl=0>.

Once the coding of is completed, we put the training data through the HK algorithm. One practical challenge that we encountered during the implementation of the HK method, however, is that ReadMe, the R package that performs the estimation procedure, has a memory constraint and therefore cannot handle datasets that are too big. To circumvent this problem, we break up the data

by quarter and conduct separate estimations on each smaller dataset. The proportions reported in the text are the averages of all quarterly estimates. We also inspected the quarterly distributions closely and found that they are quite stable over time.

To evaluate the quality of our estimates, we conducted ten-fold cross-validations. That is, we randomly split the training set into ten equally-sized sets, used each of the ten sets as the test set sequentially, and compared estimated proportions in the test set with true (hand-coded) proportions. The root mean squared error (RMSE) of the 10-fold cross validations is as low as 0.023 for the issue classification, meaning that an estimate of an issue proportion misses the true proportion by just 2.3% on average. The RMSE for location are slightly higher, at 0.057, but accounting for this uncertainty still does not change our substantive conclusion that a substantial share of the LLMB petitions are originated from rural and suburban areas.

## **F Details on Estimating Government Policy Priorities**

### **F.1 Methodology**

We use a Latent Dirichlet Allocation (LDA) model (Blei, Ng, & Jordan, 2003) to estimate the government's attention to different policy areas in the work reports. As we will show, the algorithm discovers highly substantive and coherent topics, and measurement errors in estimated topic proportions are likely to be small. The advantages of estimating topic models over human reading and coding in our case are threefold. First, the algorithm clusters words strictly according to their co-occurrence patterns, thus avoiding the arbitrariness and errors in hand coding. The somehow ambiguous boundary between topics and the fluid nature of language makes it challenging for a human coder to consistently parse the text with some pre-specified rules, where such rules may be hard to define in the first place. An LDA can not only group together words with similar semantics, but also words conforming to similar wording styles. This latter type of distinction is particularly elusive to human eyes. Second, many policies are multi-faceted, and can be attributed to more than one areas. For example, a discussion of building infrastructure is related to both the economic goal

of promoting GDP growth and the welfare goal of improving people's living conditions. In this situation it is not immediately clear how one should code this piece of discussion. LDA solves this issue by allowing both topics to give rise to the word *Infrastructure*, albeit possibly with different probabilities. Which topic a particular occurrence of *Infrastructure* belongs to is then obtained via Bayesian estimation. Last but not least, using an automated algorithm tremendously reduces the cost of parsing thousands of lengthy and dry policy documents, making them a new source of data for quantitative analysis.

One might not want to take a topic model literally. That is, the data generating process of a topic model could be very different from the way these reports are actually written. However, LDA, the most basic form of probabilistic topic models, has been shown to exhibit very good performance in a wide range of applications. Blei (2012) provides a survey of the fruitful applications of LDA in political science, psychology, population genetics, computer vision, etc. Our corpus of government work reports turns out to be a particularly good testing field for LDA, because the formal and precise wording in the reports greatly reduces noise. As a result, the topics are well-demarcated, and the majority of them are substantive (policy relevant) topics instead of "wording style" clusters.

## F.2 Estimation Procedure

We first carry out *word segmentation* on the government work reports. Chinese characters are not naturally separated from each other as in many Indo-European languages. Therefore one has to start with segmenting and demarcating the text. We apply the *Jieba* segmentation module (qinwf, 2016) to the corpus, which uses a maximum probability segmentation model and a Hidden Markov Model (HMM) to do the segmentation. The algorithm combines an existing dictionary of Chinese words and the ability to learn new words from the text. The *Jieba* package has a proven record in word segmentation, and is 5-20 times faster than other packages. We examined the segmented text, and confirmed that the segmentation quality is very good. Special terms such as *Three Represents*, *Deng Xiaoping Theory* and *18th National Congress of the Communist Party of China* are correctly identified as a single word. The segmented text is naturally tokenized, and each piece becomes

a token in subsequent computation. An unintended advantage of word segmentation is that our tokens are meaningful phrases instead of simple unigrams. Typical tokens are like *reform and open up*, *urban and rural residents* and *poverty alleviation and development*. These phrases clearly contain more information than unigrams.

We remove all punctuation, numbers and English words (such as *GDP*) from the text to focus on the Chinese vocabulary. We also remove a standard list of Chinese stop words - common words that are not really meaningful. These include words such as *some*, *both* and *why*. Because there is no inflection in Chinese, one does not need to stem the text (i.e., reducing words to their root forms). Each word in our vocabulary denotes a unique meaning. The precise and informative nature of the government reports further adds to the information contained in the preprocessed text.

Due to properties of the Dirichlet distribution, the algorithm tends to spread a topic across few words and a document across few topics. Therefore LDA is able to find topics of much higher quality when we define each paragraph of the reports as a document  $\mathbf{w}$ , instead of defining each report as a document, as a paragraph is more likely to focus on a single issue. We train LDA on reports spanning from the year 2000 to the year 2013. These reports contain 440,202 paragraphs and 3,426,528 tokens in total. The LDA is thus trained on 440,202 documents. Once the model is trained with paragraphs as documents, we infer topic proportions in whole reports with Gibbs-sampling style re-sampling.

The data generating process of LDA is as follows.

A document  $\mathbf{w} = (w_1, \dots, w_N)$  contains  $N$  words. Each word  $w_i$  is a member of the vocabulary of  $V$  words  $\{1, \dots, V\}$ . There are  $k$  topics,  $(z_1, \dots, z_k)$ .  $k$  has to be specified by the modeler. Each topic is a distribution over the words. These are characterized by a  $k \times V$  matrix  $\beta$ , where  $\beta_{ij} = p(w^j = i | z^i = 1)$ .

**Step 1:** The term distribution  $\beta$  is drawn for each topic from a Dirichlet distribution with parameter  $\delta$

$$\beta \sim \text{Dirichlet}(\delta)$$

**Step 2:** The topic proportions of document  $\mathbf{w}$  are drawn from a Dirichlet distribution with parameter  $\alpha$

$$\theta \sim \text{Dirichlet}(\alpha),$$

so that  $\theta = (\theta_1, \dots, \theta_k)$ , where  $\theta_i$  is the proportion of topic  $i$ .

**Step 3:** For each of the  $N$  words  $w_n$ ,

(a) Choose a topic  $z_n \sim \text{multinomial}(\theta)$ .

(b) Choose a word  $w_n$  from  $p(w_n|z_n, \beta)$ , a multinomial probability conditioned on the topic  $z_n$ .

One needs to specify the number of topics for the model. In the following analysis, we use results from both 10 topics and 20 topics. These are reasonable numbers of policy areas generally discussed in a report. We use the Mallet program, developed by McCallum (2002), to estimate LDA. Estimation proceeds through Gibbs sampling. We monitor Markov Chain convergence in all cases. The chains generally converge before 1,000 iterations. LDA performed very well in generating meaningful clustering of words, and some of these clusters would not be easily detected by a human coder. The vast majority of the topics consist of highly coherent words, and are easily interpretable. The complete term distribution of each topic is available upon request. Section F.4 provides the twenty highest-probability words for each topic of the 10-topic LDA, where we have named each topic. Top words for the 20-topic LDA are also available upon request.

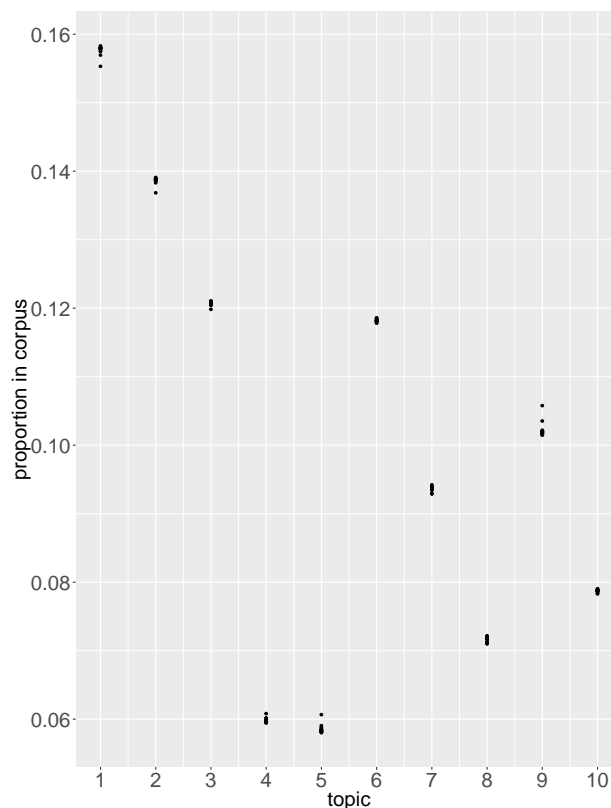
We use estimated topic proportions for each report  $\hat{\theta}_d$  as the key independent variable for our main analysis. In the context of Gibbs sampling, these are simply the proportions of words assigned to each topic in a given report in the stationary distribution of the Markov Chains.

### F.3 Post-Estimation Diagnostics

We examine the posterior distribution of the topic proportions to gauge the amount of noise in them. The theory of MCMC implies that the posterior distribution is the stationary distribution of the Markov Chain. Of the 1,000 sampling iterations, we discard the first 300 iterations as burn-in

period, and use a thinning interval of 50. That is, starting from the 300th iteration, we take one sample every 50 iterations. In each sample, we calculate the proportion of words assigned to each topic in the whole corpus. This yields fifteen data points for topic proportions in the whole corpus. Figure A.2 plots these points. The proximity of these points for each given topic is striking, to the extent that they often look like one point. In any case, the range in proportion is smaller than 0.005. This implies that the posterior distribution is very tight, and our estimated topic proportions are very precise. To the extent that the topics are meaningful encapsulations of policy, measurement errors in policy priorities are likely to be very small.

Figure A.2: Posterior Distribution of Topic Proportions in the Whole Corpus



We perform several post-estimation diagnostics to evaluate the qualities of the topics, focusing on two metrics. The first is *coherence*, which measures the tendency for top words in a topic to

appear together (Mimno, Wallach, Talley, Leenders, & McCallum, 2011). It is defined as

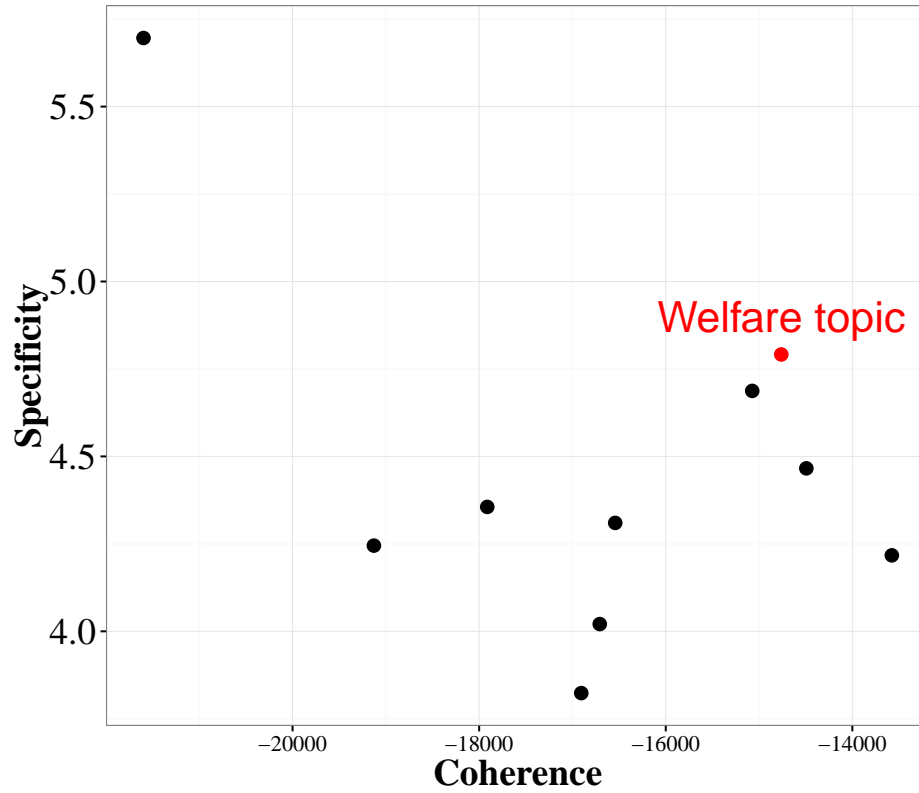
$$\sum_i \sum_{j < i} \log \frac{N(w_j, w_i) + \beta}{N(w_i)}$$

where  $w_i$  is the  $i$ th ranked word in a given topic,  $N(w_i)$  is the number of documents that contain  $w_i$ ,  $N(w_j, w_i)$  is the number of documents that contain both  $w_j$  and  $w_i$ , and  $\beta$  is a smoothing parameter. A higher coherence indicates that words in this topic are more likely to co-occur. The second one is *specificity*, which measures the distinctiveness of the words in a given topic in comparison to a uniform distribution of words, as measured by the Kullback-Leibler divergence between the word distribution of a given topic and the uniform distribution. The more specific a topic is, the greater weight it puts on a somewhat unique set of words (rather than just being a general “background” topic).

Figure A.3 displays the relative positions of the 10 estimated topics in terms of both coherence (x axis) and specificity (y axis). It turns out that the welfare topic is ranked among the highest in both dimensions. This suggests that the welfare topic is one of the best quality topics generated by the 10-topic LDA model.



Figure A.3: Topic Quality: Coherence vs. Specificity



*Note:* This figure illustrates the quality of the welfare topic in terms of both specificity and coherence, in comparison with other nine topics.

## F.4 Estimated Government Work Report Topics

### Topic 1: business attraction and industrial development

attract businesses, high technology, further, development zones, hundred million US dollars, value added, industrial parks, utilize foreign capital, open up, infrastructure, small and medium enterprises, emerging industries, competitiveness, develop vigorously, manufacturing, commercial zones, industrialization, major projects, focus projects, service industry

招商引资, 高新技术, 进一步, 开发区, 亿美元, 增加值, 工业园区, 利用外资, 对外开放, 基础设施, 中小企业, 新兴产业, 竞争力, 大力发展, 制造业, 产业园, 工业化, 重大项目, 重点项目, 服务业

### Topic 2: infrastructure and urban development

infrastructure, highway, further, overall planning, treatment plant, ten thousand square meters, project construction, environmental protection, urbanization, preliminary work, ecological environment, small cities, square kilometers, major projects, integration, comprehensive management, coverage rate, construction project, ecological improvement, polluted water treatment

基础设施, 高速公路, 进一步, 总体规划, 处理厂, 万平方米, 工程建设, 环境保护, 城镇化, 前期工作, 生态环境, 小城镇, 平方公里, 重点项目, 一体化, 综合治理, 覆盖率, 建设项目, 生态建设, 污水处理

### **Topic 3: culture, mandatory education, and public health**

further, compulsory education, family planning, spiritual advancement, public health, health care, carry out thoroughly, primary and secondary schools, broadcasting system, service system, develop vigorously, culture industry, carry out extensively, institutional reform, campaign of creating, socialism, mass-line, community health, women and children

进一步, 义务教育, 计划生育, 精神文明, 公共卫生, 医疗卫生, 深入开展, 中小学, 广播电视, 服务体系, 大力发展, 文化产业, 广泛开展, 体制改革, 创建活动, 社会主义, 群众性, 社区卫生, 妇女儿童

### **Topic 4: general economic indicators**

gross product, fixed assets, net income, disposable, consumption goods, fiscal revenue, total retail sales, urban residents, value added, economic and societal, unemployment rate, people's government, growth rate, total income, provincial government, democratic parties, percentage points, household consumption, all sectors of society, rank and file of the armed police

生产总值, 固定资产, 纯收入, 可支配, 消费品, 财政收入, 零售总额, 城镇居民, 增加值, 经济社会, 失业率, 人民政府, 增长率, 总收入, 省政府, 民主党派, 百分点, 居民消费, 各界人士, 武警官兵

### **Topic 5: real estate and high valued-added service industry**

service industry, tourism, develop vigorously, further, tertiary industry, real estate, culture industry, ten thousand people, logistics industry, total income, infrastructure, value added, real estate business, e-commerce, logistics parks, home and abroad, financial institutions, wholesale markets,

consumption goods, informatization

服务业, 旅游业, 大力发展, 进一步, 第三产业, 房地产, 文化产业, 万人次, 物流业, 总收入, 基础设施, 增加值, 房地产业, 电子商务, 物流园区, 国内外, 金融机构, 批发市场, 消费品, 信息化

#### **Topic 6: political litany**

economic and societal, views on development, further, Three Represents, implementation and solidification, Deng Xiaoping Theory, Eleventh Five-Year, socialism, moderately prosperous society, modernization, provincial government, Twelfth Five-Year, liberalize thoughts, reform and open up, industrialization, urbanization, structural adjustment, deepen the reforms, great standard-bearer, CPC Central Committee

经济社会, 发展观, 进一步, 三个代表, 贯彻落实, 邓小平理论, 十一五, 社会主义, 小康社会, 现代化, 省政府, 十二五, 解放思想, 改革开放, 工业化, 城镇化, 结构调整, 深化改革, 伟大旗帜, 党中央

#### **Topic 7: political institutions**

further, administer by law, accountability system, representatives of People's Congress, Standing Committee, democratic supervision, civil servants, carry out thoroughly, enhance clean politics, democratic parties, CPPCC members, Federation of Industry and Commerce, functions of the government, public service, people's associations, independents, party ethos and clean politics, democracy and law, public interest, establish and improve

进一步, 依法行政, 责任制, 人大代表, 常委会, 民主监督, 公务员, 深入开展, 廉政建设, 民主党派, 政协委员, 工商联, 政府职能, 公共服务, 人民团体, 无党派人士, 党风廉政, 民主法制, 群众利益, 建立健全

#### **Topic 8: agriculture modernization**

agricultural produce, commercialization, leading agricultural businesses, develop vigorously, modern agriculture, standardization, further, infrastructure, herding, labor force, organic, structural adjustment, poverty alleviation and development, agriculture and herding, demonstration zone, cooperatives, production capacity, farmers and herders, farmer income, large-scale

农产品, 产业化, 龙头企业, 大力发展, 现代农业, 标准化, 进一步, 基础设施, 畜牧业, 劳动力, 无公害, 结构调整, 扶贫开发, 农牧业, 示范区, 合作社, 生产能力, 农牧民, 农民收入, 规模化

#### **Topic 9: market reform**

further, institutional reforms, public safety, state-owned enterprises, non-public ownership, state-owned assets, public sector organizations, small and medium enterprises, crack down ruthlessly, comprehensive management, carry out thoroughly, accountability system, investment and financing, financial institutions, private firms, management system, establish and improve, letters and visits affairs, various reforms, misdemeanor and felony

进一步, 体制改革, 社会治安, 国有企业, 非公有制, 国有资产, 事业单位, 中小企业, 严厉打击, 综合治理, 深入开展, 责任制, 投融资, 金融机构, 民营企业, 管理体制, 建立健全, 信访工作, 各项改革, 违法犯罪

#### **Topic 10: social welfare**

social safety net, endowment insurance, medical insurance, urban and rural residents, urban residents, further, labor force, ten thousand square meters, social security, coverage, migrant workers, new rural, affordable, the disabled, pension, people with financial difficulties, low income, graduates, unemployment rate, urban employees

社会保障, 养老保险, 医疗保险, 城乡居民, 城镇居民, 进一步, 劳动力, 万平方米, 社会保险, 覆盖面, 农民工, 新型农村, 保障性, 残疾人, 养老金, 困难群众, 低收入, 毕业生, 失业率, 城镇职工

### **F.5 A Sample of GWR with Welfare Topics Highlighted**

The following paragraph is taken from the 2009 GWR of Luoyang city, with key words under the social welfare topic highlighted in red.

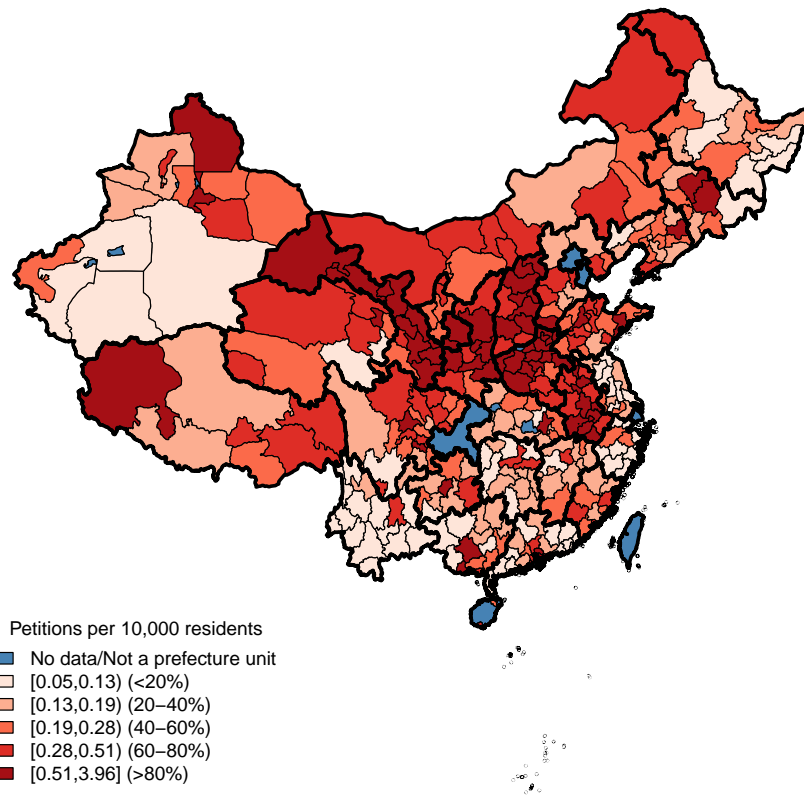
[All levels and departments should]...carefully work on the issue of employment and **reemployment**, and fully implement supporting policies such as **social security** subsidies and tax breaks, with the goal of making 200 million yuan of **microfinance loans**

in the whole year, adding 100 thousands urban jobs, and developing 1000 new **commonweal** posts...[We need to] encourage **migrant workers** to return home to set up businesses and to perfect the urban and rural **social security** system...[We also need to] strive to become a pilot city of rural **pension**, expedite the development of a **social security** system for **landless farmers**...[and] increase the construction of **affordable housing**.

## G Geographic Distribution of Online Petitions

Figure A.4 illustrates the geographic variations in the average number of petitions per 10,000 residents during the sample period. Interestingly, consistent with the pattern at the individual level, the regional variations in aggregate participation also appear to be negatively associated with economic wealth: The more developed coastal areas actually saw lower per capita participation compared to many less developed, interior regions.

Figure A.4: Average Number of Petitions per Capita, 2008-2013



## H Details on Estimating and Classifying Online Petitions using LDA Models

### H.1 Estimating Petition Topics

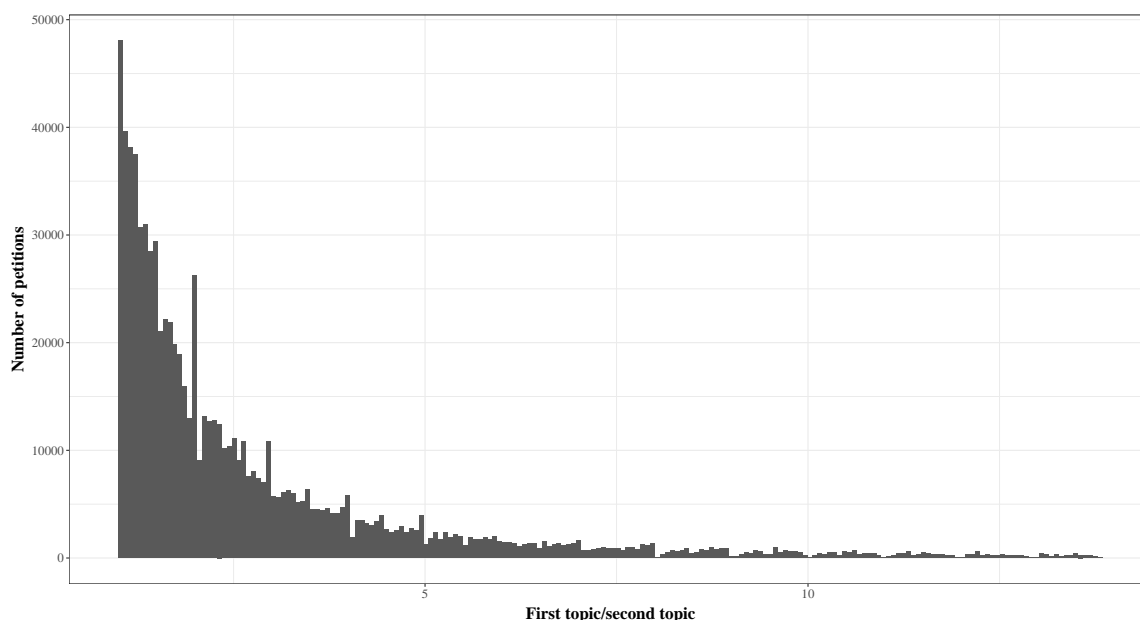
In Section H, we use readMe to infer the aggregate proportions of different issues on LLMB. However, our empirical analysis also requires classification of individual petitions, which readMe could not provide. To remedy this issue, we once again turn to LDA models. We fit 20-, 30-, and 40-topic LDA models to the corpus of petitions, treating each petition as a document. We discard words that contain just one Chinese character, as these are predominantly stop words. Section H.2 presents top words of a 30-topic LDA model. While the list of top words gives an intuitive impression that the vast majority of topics are on substantive issues (rather than specific language

usage patterns such as local slang or polite formulas) and are well-separated from each other, we also present some quantitative measures supporting the high quality of the topics. We note that these topics appear to be more meaningful than those in a typical LDA application on the English language, partly due to the succinct and formal nature of the written Chinese language.

For each topic, we calculate its “exclusivity”, which measures the extent to which top words for a topic are not top words for other topics. This is essentially a measure of how well topics are separated from each other. Formally, it is defined as the average (over 100 top words) of the ratio between probability of a word given a topic and sum of the probabilities of that word in all topics. For the 30-topic model, the minimum of 30 exclusivity scores is 0.1363, and the maximum is 0.4932. Even the minimum is much larger than the case of identical topics, in which case each exclusivity score would be  $1/30$ . This indicates that the estimated topics are highly distinct from each other and can identify different types of petitions.

Since we define each petition’s maximum topic as the petition’s topic, we want to confirm that petitions do have dominant topics that indicate their types. Figure A.5 shows the ratio between largest topic proportion and second-largest topic proportion for all petitions. It can be seen that this ratio is generally large, and has a long right tail. The mean of the distribution is 5.55, and the median is 2.00. This confirms that the largest topic in a petition does capture the main issue being discussed.

Figure A.5: Clear Dominant Topic in Petitions



Note: This figure plots the distribution of largest topic proportion over second-largest topic proportion in petitions. The 5% right tail is truncated.

## H.2 Details about Topic Classification

### *Rural Pocketbook (RP)*

Topic 16: villager, road, path, leader, secretary, now, have not, build a road, travel, cement road, suddenly, road surface, bumpy, raining day

村民, 公路, 道路, 领导, 书记, 现在, 没有, 希望, 修路, 出行, 水泥路, 一下, 路面, 坑坑洼洼, 下雨天

Topic 23: village, farmer, have not, secretary, leader, policy, state, compensation, subsidy, low-income allowance, this year, loan, one, finance, poverty alleviation

农村, 农民, 没有, 书记, 领导, 政策, 国家, 补贴, 补助, 低保, 今年, 贷款, 一个, 资金, 扶贫

Topic 29: villager, land, have not, secretary, farmland, compensation, land taking, village committee, village cadres, occupy, my home, rural residential land, forcefully, compensation package

村民, 土地, 没有, 书记, 耕地, 领导, 补偿, 征地, 村委会, 村干部, 占用, 我家, 宅基地, 强行, 补偿款



### ***Urban Pocketbook (UP)***

Topic 7: hospitals, father, reimbursement, doctors, life, mother, treatment, family, medical insurance, children, cannot, handicapped, elderly, in-patient, expenses

医院, 父亲, 报销, 医生, 生活, 母亲, 治疗, 家庭, 医保, 孩子, 不能, 残疾人, 老人, 住院, 费用

Topic 8: demolishing, home, settlement, upgrade, government, house, have not, compensation, leader, my home, construction, planning, secretary, shanty-town, relocated households

拆迁, 房子, 安置, 改造, 政府, 房屋, 没有, 补偿, 领导, 我家, 建设, 规划, 书记, 棚户区, 拆迁户

Topic 11: household registration, children, handle, have not, policy, certificate, work, cannot, police branch, one, parents, leader, need, please, residence

户口, 孩子, 办理, 没有, 政策, 证明, 工作, 不能, 派出所, 一个, 父母, 领导, 需要, 请问, 户籍

Topic 22: company, salary, employee, firm, have not, leader, migrant worker, limited-liability company, worker, secretary, unit, arrears, projects, labor, ten thousand yuan

公司, 工资, 职工, 企业, 没有, 领导, 农民工, 有限公司, 工人, 书记, 单位, 拖欠, 工程, 劳动, 万元

Topic 26: salary, work, teacher, have not, unit, personnel, retirement, employee, compensation package, life, secretary, policy, leader, state, now

工资, 工作, 教师, 没有, 单位, 人员, 退休, 职工, 待遇, 生活, 书记, 政策, 领导, 国家, 现在

### ***Rural Non-Pocketbook (RN)***

Topic 13: villager, farmer, reservoir, severe, cultivation, cause, damage, production, government, extraction, river course, now, land, secretary, farmland

村民, 农民, 水库, 严重, 种植, 造成, 破坏, 生产, 政府, 开采, 河道, 现在, 土地, 书记, 农田

### ***Urban Non-Pocketbook (UN)***

Topic 1: residential compound, property owner, real estate property, have not, residents, problem, elevator, occupants, property management company, inside, management, property manage-

ment fee, garden, agency, community

小区, 业主, 物业, 没有, 居民, 问题, 电梯, 住户, 物业公司, 区内, 管理, 物业费, 花园, 部门, 社区

Topic 2: vehicle, road, traffic, severe, safety, passenger, cross-road, streetlight, do not have, agency, segment of road, travel, cause, influence, hope

车辆, 道路, 交通, 严重, 安全, 行人, 路口, 路灯, 没有, 部门, 路段, 出行, 造成, 影响, 希望

Topic 3: secretary, problem, solve, leader, hope, hello, not yet, respect, reflect, mayor, thank you, take time from a busy schedule, attention, now, ask

书记, 问题, 解决, 领导, 希望, 您好, 没有, 尊敬, 反映, 市长, 谢谢, 百忙之中, 关注, 现在, 过问

Topic 4: garbage, pollution, severe, residents, environment, life, waste water, influence, one, nearby, hope, leader, health, emission, production

垃圾, 污染, 严重, 居民, 环境, 生活, 污水, 影响, 一个, 附近, 希望, 领导, 健康, 排放, 生产

Topic 5: bus, public transport, driver, taxi, vehicle, passenger, convenience, time, have not, travel, hours, hope, traffic, car, train station

公交车, 公交, 司机, 出租车, 车辆, 乘客, 方便, 时间, 没有, 出行, 小时, 希望, 交通, 汽车, 火车站

Topic 6: police station, one, pyramid selling scam, police, personnel, happen, police department, at that time, call the police, tour guide, touring, have not, hope, friends, police

派出所, 一个, 传销, 警察, 人员, 发生, 公安局, 当时, 报警, 导游, 旅游, 没有, 希望, 朋友, 民警

Topic 9: have not, now, one, know, leader, folks, hope, really, secretary, government, once, why not, cannot, location, see

没有, 现在, 一个, 知道, 领导, 老百姓, 希望, 真的, 书记, 政府, 一下, 难道, 不能, 地方, 看到

Topic 10: planning, construction, railway, residents, convenience, wide road, traffic, highway, have not, public transport, residential compound, connect, please, nearby

规划, 建设, 铁路, 居民, 方便, 出行, 大道, 交通, 高速, 没有, 公交, 小区, 开通, 请问, 附近

Topic 12: house, residents, severe, safety, construction, department, remove, in the process of construction, occupants, house, problem, influence, residential compounds, have not, cause

房屋, 居民, 严重, 安全, 建筑, 部门, 拆除, 施工, 住户, 房子, 问题, 影响, 小区, 没有, 造成

Topic 14: operate, market, department, one, influence, severe, urban management officer, hope, manage, environment, Internet cafe, leader, secretary, gambling, commercial tenant

经营, 市场, 部门, 一个, 影响, 严重, 城管, 希望, 管理, 环境, 网吧, 领导, 书记, 赌博, 商户

Topic 15: school, students, children, teacher, primary school, parents, kindergarten, education, middle school, make-up class, attend school, education bureau, leader, learn, one

学校, 学生, 孩子, 老师, 小学, 家长, 幼儿园, 教育, 中学, 补课, 上学, 教育局, 领导, 学习, 一个

Topic 17: developer, property owner, real estate transaction, apartment, have not, now, residential compound, property ownership certificate, handle, contract, government, develop, leader, already, purchase

开发商, 业主, 交房, 房子, 没有, 现在, 小区, 房产证, 办理, 合同, 政府, 开发, 领导, 已经, 购买

Topic 18: test, work, college students, teacher, driving school, civil servants, have not, not yet, graduation, participate, professional, village officer, employment, one, graduating student, test taker

考试, 工作, 大学生, 教师, 驾校, 公务员, 没有, 毕业, 参加, 专业, 村官, 就业, 一名, 毕业生, 考生

Topic 19: regulation, relevant, department, undertake, require, law, illegal, condition, government, court, report, have not, state, behavior, unit

规定, 相关, 部门, 进行, 要求, 法律, 违法, 情况, 政府, 法院, 反映, 没有, 国家, 行为, 单位

Topic 20: leader, local, people's daily website, message board, source, secretary, hello, once, respect, hope, have not, hi, now, thank you, pay attention to

领导, 地方, 人民网, 留言板, 来源, 书记, 您好, 一下, 尊敬, 希望, 没有, 你好, 现在, 谢谢,

关注

Topic 21: charge fees, charge, fee, regulation, standard, price, state, whether, reasonable, unreasonable charges, please, document, natural gas, request

收费, 收取, 费用, 规定, 标准, 价格, 国家, 是否, 合理, 乱收费, 缴纳, 请问, 文件, 天然气, 要求

Topic 24: telephone, have not, handle, staff, information, Internet, company, cannot, one, broadband, make a call, complain, bank, mobile, phone

电话, 没有, 办理, 工作人员, 信息, 网络, 公司, 不能, 一个, 宽带, 打电话, 投诉, 银行, 移动, 手机

Topic 25: government, problem, mass, work, people, society, leader, folks, cadre, department, one, real, hope, proceed, should

政府, 问题, 群众, 工作, 人民, 社会, 领导, 百姓, 干部, 部门, 一个, 真正, 希望, 进行, 应该

Topic 27: develop, construction, city, economy, one, suggestion, travel, culture, hope, rural areas, people, whole country, hometown, environment, secretary

发展, 建设, 城市, 经济, 一个, 建议, 旅游, 文化, 希望, 农村, 人民, 全国, 家乡, 环境, 书记

Topic 28: residents, noise, affect, severe, residential compounds, at night, life, disruptive, department, everyday, in the process of construction, normal, environment, one

居民, 噪音, 影响, 严重, 小区, 休息, 晚上, 生活, 扰民, 部门, 每天, 施工, 正常, 环境, 一个

Topic 30: residential compounds, heating, company, solve, residents, have not, heat, water supply, heat supply, occupants, leader, life, now, shutdown water supply

小区, 供暖, 公司, 解决, 居民, 问题, 没有, 暖气, 自来水, 供热, 住户, 领导, 生活, 现在, 停水

## I Data on Collective Protests

The Collective Incidents Dataset (CID) is a dataset compiled by the institute of sociology at the Chinese Academy of Social Sciences (CASS), one of the major state-sponsored think tanks in

China. The project was motivated by the government's concern about the rising social instability since 2000. Between 2009 and 2013, the project collected on a daily basis information about collective protests and other major mass incidents,<sup>28</sup> and constructed a GIS based dataset, which allows researchers to analyze the spatial distribution and diffusion of mass incidents. Information about protests came from a variety of sources, including internal government documents, state-owned and commercial newspapers, TV reports, websites and social media. Extensive search was conducted on both domestic and foreign sources to minimize potential sampling bias. The dataset contains a rich set of features regarding protests recorded, such as the course of the event, location, actors, and interaction between actors, causes and consequences, process, number of participants, duration, and the information sources. While this dataset obviously does not exhaust all protests that happened in China during this period, it does provide a decent coverage on the large-scale protests, which are most likely to confound our results by inducing substantive policy changes. For the period of interest (2008-2013), the dataset contains 841 protests, with over 85% of them involving more than a thousand participants. The project was terminated in December 2013.

The other important data source is the *China Strikes* website, which records instances of labor unrest in China based on a variety of sources, including reports from website visitors. This website is dedicated to covering only contentious, collective actions by *workers over workplace issues*. Although this dataset is limited in the scope of protests covered, one of its key advantage is that it is maintained by researchers outside mainland China<sup>29</sup> and therefore less susceptible to censorship issues. For each entry, information about the date, time, location, issues, and identity of participants are provided. Sometimes, an entry will also contain photos and video clips that verify the occurrence of a protest. Between 2008 and 2012,<sup>30</sup> this dataset contains information for a total of 721 incidences.

We combine all the unique records from both sources into a new dataset, which contains a total of 1268 protests between 2008 and 2013. Most notably, even though these two datasets

---

<sup>28</sup>Retrospective collection was also conducted on protests that took place prior to the project start date.

<sup>29</sup>Currently, this website is maintained by Manfred Elfstrom, a postdoctoral fellow at the Kennedy School of Harvard University.

<sup>30</sup>The latest data available on the website are at December 2012.

are supposedly collected by different teams and through different methods, we find a correlation of 0.74 between the two in terms of aggregate number of protests at the city level. This large correlation gives us confidence that there is a common underlying pattern of social instability that both datasets have managed to capture. In the combined dataset, a city experiences about 0.7 protests per year with a large standard deviation of 2.4. A simple correlation test suggests there is indeed a positive and significant association between the frequency of offline protest and the volume of online petitions ( $\rho = 0.23$ ,  $p < 0.0001$ ).

## **J Validating the Link between Petition and Rural Participants**

One potential concern with this procedure is whether our classification of LDA topics can accurately reflect participants' backgrounds. In this section, we conduct several additional tests to assess the validity of our LDA classifications. Specifically, we hypothesize that if rural petitions are indeed filed by those residing in the countryside, the volume of such petitions should be positively correlated with the size of the rural sector. To test this, we regress rural petition volumes on two measures of local rural sector size—the size of rural population and the share of rural area in the whole city. The results are shown in [A.4](#). The first two columns show that total petitions are positively correlated with both rural population size and the share of rural area. This is consistent with our claim that the LLMB is a tool that is more frequently used by rural citizens. Columns 3 through 6 further show that rural sector size is strongly and positively associated with the volumes of both total rural petitions and rural pocketbook petitions more specifically. By contrast, urban population size has no relationship with these two types of petitions. Overall, these patterns suggest that rural petitions are indeed more common in localities with a larger rural sector and support the empirical validity of our classification decisions.

Table A.4: Validating the Link between Online Petitions and Rural Participants

|   | Total Petitions<br>(Log) |                       | Rural Petitions<br>(Log) |                       | Rural Pocketbook<br>Petitions (Log) |                       |
|---|--------------------------|-----------------------|--------------------------|-----------------------|-------------------------------------|-----------------------|
|   | (1)                      | (2)                   | (3)                      | (4)                   | (5)                                 | (6)                   |
| Log rural population (2005 population survey) | 0.3383***<br>(0.0576)    |                       | 0.6279***<br>(0.0578)    |                       | 0.6388***<br>(0.0586)               |                       |
| Log urban population (2005 population survey) | 0.0845<br>(0.1052)       |                       | -0.0876<br>(0.1116)      |                       | -0.1053<br>(0.1131)                 |                       |
| % of rural area                               |                          | 2.1245*<br>(1.1899)   |                          | 8.9996***<br>(2.4714) |                                     | 8.6351***<br>(2.4014) |
| Log GDP                                       | 0.3559***<br>(0.0871)    | 0.5236***<br>(0.0656) | 0.1313<br>(0.0899)       | 0.3711***<br>(0.0703) | 0.1450<br>(0.0907)                  | 0.3749***<br>(0.0708) |
| Year fixed-effects                            | ✓                        | ✓                     | ✓                        | ✓                     | ✓                                   | ✓                     |
| R <sup>2</sup>                                | 0.37                     | 0.26                  | 0.42                     | 0.25                  | 0.41                                | 0.23                  |
| Number of Cities                              | 304                      | 281                   | 304                      | 281                   | 304                                 | 281                   |
| Observations                                  | 1821                     | 1652                  | 1821                     | 1652                  | 1821                                | 1652                  |

**Note:** This table shows that the volumes of both overall petitions and rural-specific petitions are positively correlated with the size of the rural sector in a locality. We use two different measures for rural sector size: rural population and rural area share. Robust standard errors clustered at city level are reported in parentheses.

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$  (two-tailed test)

## K Baseline Results with All Coefficients

Table A.5: Baseline Results with Coefficients for All Control Variables

|   | $\Delta$ Welfare Topic at $t + 1$ (10-topic) |                       |                       |                       |                       |
|---|--|-----------------------|-----------------------|-----------------------|-----------------------|
|   | (1)  | (2)                   | (3)                   | (4)                   | (5)                   |
| Log petitions                           | 0.0048**<br>(0.0021)                         |                       |                       |                       |                       |
| Log petitions (pocketbook)              |  | 0.0085***<br>(0.0029) |                       |                       |                       |
| Log petitions (non-pocketbook)          |  | -0.0025<br>(0.0028)   |                       |                       |                       |
| Log petitions (rural pocketbook)        |  |                       | 0.0073***<br>(0.0021) | 0.0075***<br>(0.0022) | 0.0076***<br>(0.0022) |
| Log petitions (urban pocketbook)        |  |                       | 0.0019<br>(0.0022)    | 0.0022<br>(0.0023)    | 0.0021<br>(0.0023)    |
| Log petitions (rural non-pocketbook)    |  |                       | -0.0005<br>(0.0016)   | -0.0004<br>(0.0016)   | -0.0005<br>(0.0016)   |
| Log petitions (urban non-pocketbook)    |  |                       | -0.0023<br>(0.0028)   | -0.0026<br>(0.0028)   | -0.0026<br>(0.0029)   |
| Log protests                            |  |                       |                       | -0.0046**<br>(0.0023) | -0.0047**<br>(0.0023) |
| Log employment (10,000 persons)         |  |                       |                       | 0.0063<br>(0.0044)    | 0.0063<br>(0.0044)    |
| Log GDP                                 |  |                       |                       | 0.0004<br>(0.0100)    | 0.0011<br>(0.0102)    |
| Log population                          |  |                       |                       | -0.0178**<br>(0.0083) | -0.0170**<br>(0.0082) |
| GDP growth rate                         |  |                       |                       | -0.0006**<br>(0.0003) | -0.0006**<br>(0.0003) |
| Log fiscal expenditure                  |  |                       |                       | 0.0031<br>(0.0099)    | 0.0025<br>(0.0100)    |
| Log fiscal revenue                      |  |                       |                       | -0.0061<br>(0.0065)   | -0.0060<br>(0.0067)   |
| Mayor's age                             |  |                       |                       |                       | 0.0000<br>(0.0003)    |
| City secretary's age                    |  |                       |                       |                       | 0.0002<br>(0.0003)    |
| City secretary's tenure                 |  |                       |                       |                       | -0.0001<br>(0.0005)   |
| Mayor's tenure                          |  |                       |                       |                       | -0.0009<br>(0.0006)   |
| City secretary's years of local service |  |                       |                       |                       | -0.0002<br>(0.0001)   |
| Mayor's years of local service          |  |                       |                       |                       | 0.0001<br>(0.0002)    |
| Connected city leader                   |  |                       |                       |                       | 0.0000<br>(0.0022)    |
| City and year fixed-effects             | ✓  | ✓                     | ✓                     | ✓                     | ✓                     |
| Province-specific trends                | ✓  | ✓                     | ✓                     | ✓                     | ✓                     |
| R <sup>2</sup>                          | 0.02   | 0.03                  | 0.03                  | 0.04                  | 0.04                  |
| Number of Cities                        | 299  | 299                   | 299                   | 298                   | 297                   |
| Observations                            | 1656   | 1656                  | 1656                  | 1625                  | 1624                  |

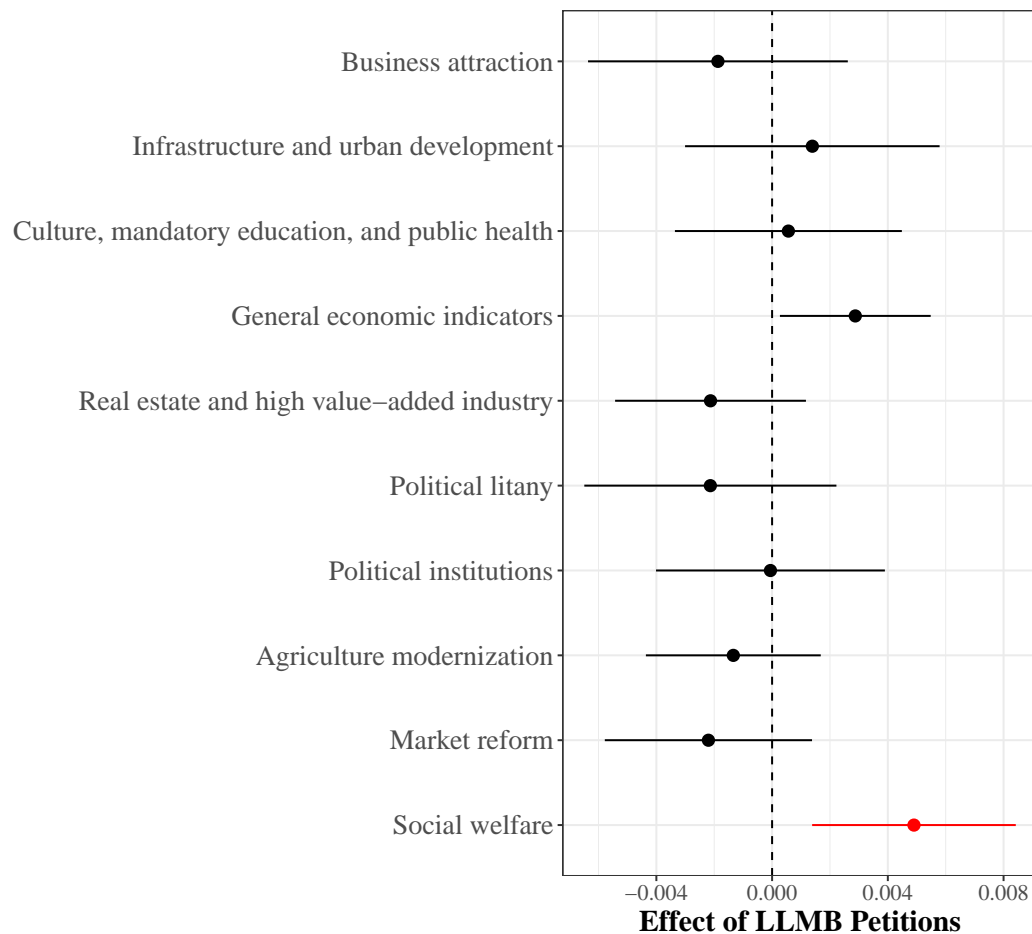
**Note:** This table shows the baseline OLS results (Table 2) with coefficient estimates for all the controls. Robust standard errors clustered at city level are reported in parentheses.

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$  (two-tailed test)



## L Effect of Petition on All GWR Policy Topics

Figure A.6: Effect of Petitions on All GWR Topics



*Note:* This figure visually presents the OLS results from regressing LLMB petitions on all ten GWR topics. The horizontal lines represent the 90% confident intervals. The model is based on Column 1 of Table 2.

## M Robustness Checks

### M.1 Baseline Estimates without City Fixed-Effects

Table A.6: Baseline Estimates without Fixed-Effects

|                                      | $\Delta$ Welfare Topic at $t + 1$ (10-topic) |                      |                      |                       |                       |
|--------------------------------------|--|----------------------|----------------------|-----------------------|-----------------------|
|                                      | (1)  | (2)                  | (3)                  | (4)                   | (5)                   |
| Log petitions                        | 0.0007*<br>(0.0004)                          |                      |                      |                       |                       |
| Log petitions (pocketbook)           |  | 0.0030**<br>(0.0015) |                      |                       |                       |
| Log petitions (non-pocketbook)       |  | -0.0021<br>(0.0014)  |                      |                       |                       |
| Log petitions (rural pocketbook)     |  |                      | 0.0021**<br>(0.0010) | 0.0038***<br>(0.0014) | 0.0039***<br>(0.0014) |
| Log petitions (urban pocketbook)     |  |                      | 0.0009<br>(0.0013)   | 0.0008<br>(0.0015)    | 0.0007<br>(0.0015)    |
| Log petitions (rural non-pocketbook) |  |                      | -0.0006<br>(0.0010)  | -0.0004<br>(0.0011)   | -0.0006<br>(0.0011)   |
| Log petitions (urban non-pocketbook) |  |                      | -0.0016<br>(0.0014)  | -0.0019<br>(0.0015)   | -0.0019<br>(0.0015)   |
| Year fixed-effects                   | ✓  | ✓                    | ✓                    | ✓                     | ✓                     |
| Economic controls                    |  |                      |                      | ✓                     | ✓                     |
| Leadership controls                  |  |                      |                      |                       | ✓                     |
| R <sup>2</sup>                       | 0.01   | 0.01                 | 0.01                 | 0.02                  | 0.02                  |
| Number of Cities                     | 299  | 299                  | 299                  | 298                   | 297                   |
| Observations                         | 1656   | 1656                 | 1656                 | 1625                  | 1624                  |

**Note:** This table shows regression results that do not include city or province fixed effects. The specifications are otherwise identical to those in Table 2.

Robust standard errors clustered at city level are reported in parentheses.

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$  (two-tailed test)

## M.2 Alternative Dependent Variable: GWR Welfare Topic from a 20-Topic Model

Table A.7: Using Welfare Topic from 20-Topic LDA Model as the Dependent Variable

|                                      | $\Delta$ Welfare Topic at $t + 1$ (20-topic) |                       |                       |                       |                       |
|--------------------------------------|--|-----------------------|-----------------------|-----------------------|-----------------------|
|                                      | (1)  | (2)                   | (3)                   | (4)                   | (5)                   |
| Log petitions                        | 0.0049***<br>(0.0019)                        |                       |                       |                       |                       |
| Log petitions (pocketbook)           |  | 0.0085***<br>(0.0025) |                       |                       |                       |
| Log petitions (non-pocketbook)       |  | -0.0024<br>(0.0025)   |                       |                       |                       |
| Log petitions (rural pocketbook)     |  |                       | 0.0067***<br>(0.0018) | 0.0069***<br>(0.0019) | 0.0070***<br>(0.0019) |
| Log petitions (urban pocketbook)     |  |                       | 0.0025<br>(0.0020)    | 0.0029<br>(0.0021)    | 0.0028<br>(0.0021)    |
| Log petitions (rural non-pocketbook) |  |                       | -0.0002<br>(0.0014)   | -0.0002<br>(0.0015)   | -0.0002<br>(0.0015)   |
| Log petitions (urban non-pocketbook) |  |                       | -0.0024<br>(0.0025)   | -0.0027<br>(0.0025)   | -0.0026<br>(0.0026)   |
| City and year fixed-effects          | ✓  | ✓                     | ✓                     | ✓                     | ✓                     |
| Province-specific trends             | ✓  | ✓                     | ✓                     | ✓                     | ✓                     |
| Socioeconomic controls               |  |                       |                       | ✓                     | ✓                     |
| Leadership controls                  |  |                       |                       |                       | ✓                     |
| R <sup>2</sup>                       | 0.03   | 0.03                  | 0.04                  | 0.04                  | 0.04                  |
| Number of Cities                     | 299  | 299                   | 299                   | 298                   | 297                   |
| Observations                         | 1656   | 1656                  | 1656                  | 1625                  | 1624                  |

**Note:** This table shows the results from OLS regressions. The dependent variables are incremental increase in the share of social welfare topic in government work reports (from 20-topic LDA model). The specification is based on Table 2. Robust standard errors clustered at city level are reported in parentheses.

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$  (two-tailed test)

## M.3 Estimation on Different Samples

To ensure the robustness of our main findings, we re-estimate the baseline regression using several different samples. First, we add back the western provinces to see if doing so would significantly change the result. As shown in Table A.8, some estimates become a bit noisier but the main patterns are substantively unchanged. Next, we split the sample by three main geographical regions: east, central, and west. The first three columns of Table A.9 show that the coefficients are roughly

comparable across different regions: they are most precisely estimated in east and central regions, but also have a large, albeit noisy estimate, for the western regions. Finally, we check if our results are driven by a small number of cities with special political status. In Column 4, we exclude the provincial capitals, which are large and politically important cities. In Column 5, we exclude cities with large presence of ethnic minorities (i.e., the autonomous cities). In both cases, our main OLS results remain positive and statistically significant.

Table A.8: Sub-Sample Analysis: Including Low Petition Provinces

|                                      | $\Delta$ Welfare Topic at $t + 1$ |                       |
|--------------------------------------|-----------------------------------|-----------------------|
|                                      | (1)                               | (2)                   |
| Log petitions                        | 0.0045**<br>(0.0020)              |                       |
| Log petitions (rural pocketbook)     |                                   | 0.0075***<br>(0.0021) |
| Log petitions (urban pocketbook)     |                                   | 0.0022<br>(0.0023)    |
| Log petitions (rural non-pocketbook) |                                   | -0.0006<br>(0.0016)   |
| Log petitions (urban non-pocketbook) |                                   | -0.0028<br>(0.0028)   |
| City and year fixed-effects          | ✓                                 | ✓                     |
| Number of Cities                     | 308                               | 308                   |
| Observations                         | 1655                              | 1655                  |

**Note:** The table reports OLS results using a sample that includes provinces with low LLMB usage (Xinjiang, Qinghai, Tibet). Robust standard errors clustered at city level are reported in parentheses.

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$  (two-tailed test)

Table A.9: Sub-Sample Analysis: By Region

|                                      | Subsample: Region    |                      |                      |   |   |
|--------------------------------------|----------------------|----------------------|----------------------|---|---|
|                                      | (1)<br>East          | (2)<br>Central       | (3)<br>West          | (4)<br>Exclude<br>Provincial<br>Capital | (5)<br>Exclude<br>Autonomous<br>City<br>(zhou/meng) |
| Log petitions (rural pocketbook)     | 0.0085**<br>(0.0042) | 0.0065**<br>(0.0025) | 0.0125<br>(0.0256)   | 0.0075***<br>(0.0022)                   | 0.0074***<br>(0.0023)                               |
| Log petitions (urban pocketbook)     | -0.0015<br>(0.0048)  | 0.0034<br>(0.0026)   | -0.0061<br>(0.0209)  | 0.0018<br>(0.0024)                      | 0.0017<br>(0.0024)                                  |
| Log petitions (rural non-pocketbook) | -0.0031<br>(0.0027)  | 0.0011<br>(0.0020)   | -0.0311*<br>(0.0155) | -0.0006<br>(0.0017)                     | -0.0006<br>(0.0017)                                 |
| Log petitions (urban non-pocketbook) | -0.0085<br>(0.0053)  | -0.0014<br>(0.0034)  | 0.0107<br>(0.0178)   | -0.0023<br>(0.0029)                     | -0.0029<br>(0.0029)                                 |
| Year fixed-effects                   | ✓                    | ✓                    | ✓                    | ✓                                       | ✓   |
| Province-specific trends             | ✓                    | ✓                    | ✓                    | ✓                                       | ✓   |
| R <sup>2</sup>                       | 0.07                 | 0.04                 | 0.73                 | 0.05                                    | 0.04  |
| Number of Cities                     | 84                   | 208                  | 16                   | 282                                     | 281   |
| Observations                         | 492                  | 1108                 | 55                   | 1514                                    | 1544  |

**Note:** This table shows regression results from subsample analysis. The results suggest that the level of policy responsiveness is comparable both across geographical regions.

Robust standard errors clustered at city level are reported in parentheses.

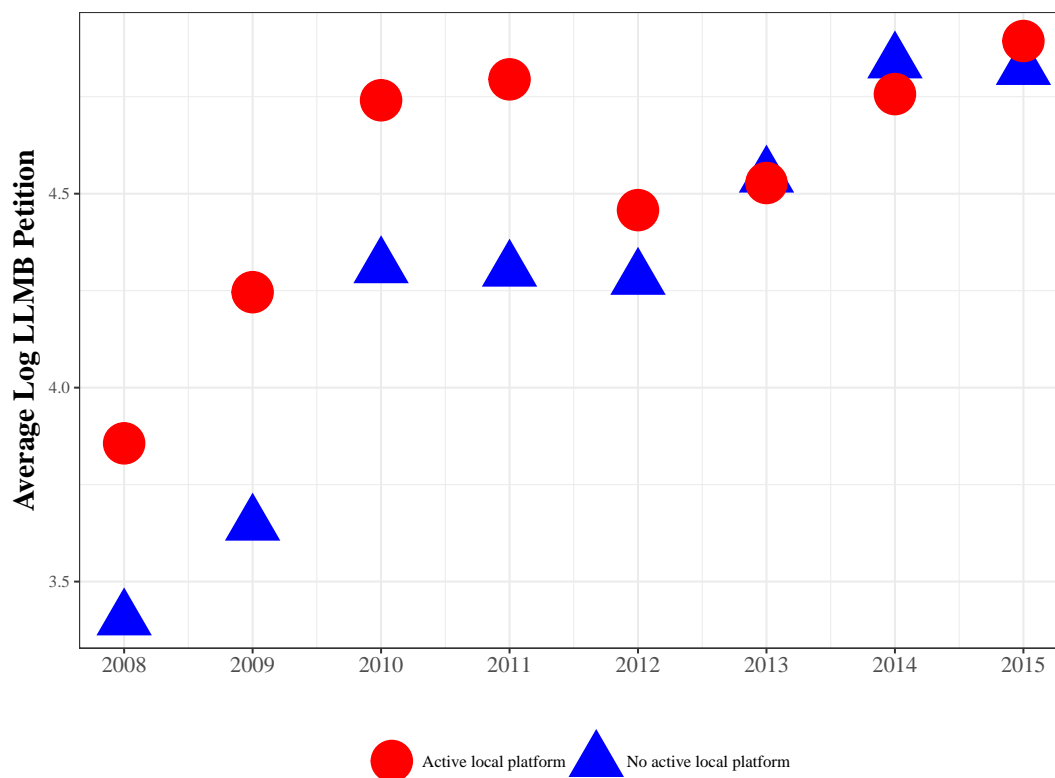
\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$  (two-tailed test)

## M.4 Addressing the Presence of Local Petition Platforms

Another potential concern one may have is that there are some localities that have developed their own petition platforms to receive citizen petitions and our analyses, which focus only on the centrally operated petition platform, do not capture petitions filed at those more localized sites. Although we do not see any reason as to why the patterns of participation in those more localized sites should be different from that in the LLMB (as citizens who engage in online petitions are likely to know and use both), our results might be affected if citizens are strategically choosing to send certain type of petitions *only* to central or local platforms. While the amount of resources and time needed for collecting systematic data on all local platforms is well beyond our current capacity, we try to address this issue by first identifying the localities that have the most active local platforms and then compare the patterns of online participation there with the rest of the country. In particular, we note that a number of local petition platforms have been awarded the China News

Prize (中国新闻奖) (the highest prize for journalistic/media work in China) because of their high popularity among local citizens. We compiled a full list of cities that have received such prizes as the “active local platform” group,<sup>31</sup> and compare LLMB participation in this group with that in other cities (i.e., no active local platform). Figure A.7 displays the annual changes in average log petitions. We can see that temporal trends of the two groups track quite closely with each other, although those with active local platforms tend to have more LLMB users during the first several years following LLMB’s launch. A simple regression analysis shows that active local platform is associated with about 35% increase in LLMB petitions between 2008 and 2013. This pattern is thus consistent with our assumption that the level of LLMB usage is representative of the level of usage in other online petition platforms.

Figure A.7: Level of LLMB Participation by Presence of Active Local Petition Platform



Second, we also conduct a subsample analysis to see if the patterns of responsiveness differ

<sup>31</sup>The list includes the following localities: (provinces) Henan, Guangdong, Liaoning, (cities) Yantai, Ningbo, Zhenjiang, and Huizhou.

between localities with and without active local platforms. The results are presented in Table A.9. We can see that the estimates are largely comparable between the two samples and the main estimate of interest is positive and highly significant even when we exclude those localities with more active local online platforms. Taken together, these results suggest that our findings are unlikely to be driven by the presence or absence of alternative local petition platforms.

Table A.10: Sub-Sample Analysis: By Presence of Local Platforms

|                                      | Subsample: Pre-existing Local Platforms |                       |
|--------------------------------------|---|-----------------------|
|                                      | (1)<br>Yes                              | (2)<br>No             |
| Log petitions (rural pocketbook)     | 0.0072<br>(0.0048)                      | 0.0072***<br>(0.0026) |
| Log petitions (urban pocketbook)     | 0.0007<br>(0.0053)                      | 0.0029<br>(0.0026)    |
| Log petitions (rural non-pocketbook) | 0.0014<br>(0.0028)                      | -0.0011<br>(0.0020)   |
| Log petitions (urban non-pocketbook) | -0.0017<br>(0.0055)                     | -0.0032<br>(0.0035)   |
| Year fixed-effects                   | ✓                                       | ✓                     |
| Province-specific trends             | ✓                                       | ✓                     |
| Adjusted R <sup>2</sup>              | 0.09                                    | 0.04                  |
| Number of Cities                     | 69                                      | 228                   |
| Observations                         | 400                                     | 1224                  |

**Note:** This table shows regression results from subsample analysis. The results suggest that the level of policy responsiveness is comparable between localities with and without large, active local petition platforms.

Robust standard errors clustered at city level are reported in parentheses.

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$  (two-tailed test)

## N Granger-Style Tests on Reverse Causality

We conduct several Granger-style tests to see if our results are driven by reverse causality. We run two additional regressions with the timing of the dependent variable set at  $t$  and  $t - 1$ , respectively. If reverse causality exists, then we would expect content of GWRs to be correlated with petition intensity in current and subsequent years. However, results presented in Table A.11 suggest that this is not the case: neither the current or the previous year's GWR welfare topic appears to be significantly correlated with the current year's volume of online petition. In the Online Appendix,

we conduct several more extensive checks by including leads and lags of welfare topic share at both the city and the provincial level, and the main results remain largely unchanged (Table A.12).

Table A.11: Alternate the Timing of the Dependent Variable

|                                  | $\Delta$ Welfare Topic (10-topic) at ... |                     |                    |                    |
|----------------------------------|--|---------------------|--------------------|--------------------|
|                                  | (1)<br>$t$                               | (2)<br>$t$          | (3)<br>$t - 1$     | (4)<br>$t - 1$     |
| Log petitions                    | -0.0005<br>(0.0022)                      |                     | 0.0020<br>(0.0025) |                    |
| Log petitions (rural pocketbook) |  | -0.0036<br>(0.0023) |                    | 0.0028<br>(0.0024) |
| City and year fixed-effects      | ✓  | ✓                   | ✓                  | ✓                  |
| Province-specific trends         | ✓  | ✓                   | ✓                  | ✓                  |
| R <sup>2</sup>                   | 0.08                                     | 0.08                | 0.09               | 0.09               |
| Number of Cities                 | 308                                      | 308                 | 300                | 300                |
| Observations                     | 1607                                     | 1607                | 1537               | 1537               |

**Note:** This table shows results from altering the timing of the dependent variable. The results suggest that earlier welfare topic proportions in GWRs is not significantly associated with subsequent petition intensity.

Robust standard errors clustered at city level are reported in parentheses.

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$  (two-tailed test)

In addition to the analyses presented in Table A.11, we also estimate several regressions that include both the current and past welfare topic shares for provincial as well as city governments. Due to the well-known Nickell bias (Nickell, 1981), we do not include city fixed-effects in these regressions to avoid the incidental parameter problem. The results are presented in Table A.12. The first two columns include lagged welfare topic share for city GWR ( $t$  and  $t - 1$ ), and Columns 3 through 5 include the welfare topic shares in provincial GWRs at  $t + 1$ ,  $t$ , and  $t - 1$ , respectively. Finally, the last column presents a model that includes all the preceding controls. Through all the models, we see that the coefficient estimate for type RP petitions remains quite stable. This lends us further confidence that the observed effect of online participation is unlikely to be driven solely by changes in local government policies.



Table A.12: Controlling for Welfare Topic Share at the Higher Level or from Earlier Periods

|   | $\Delta$ Welfare Topic at $t + 1$ (10-topic) |                        |                       |                       |                       |                        |
|---|--|------------------------|-----------------------|-----------------------|-----------------------|------------------------|
|   | (1)  | (2)                    | (3)                   | (4)                   | (5)                   | (6)                    |
| Log petitions (rural pocketbook)                | 0.0043***<br>(0.0013)                        | 0.0053***<br>(0.0017)  | 0.0055***<br>(0.0016) | 0.0076***<br>(0.0022) | 0.0054***<br>(0.0016) | 0.0039***<br>(0.0013)  |
| Log petitions (urban pocketbook)                | 0.0015<br>(0.0015)                           | 0.0012<br>(0.0018)     | 0.0014<br>(0.0018)    | 0.0022<br>(0.0023)    | 0.0014<br>(0.0018)    | 0.0016<br>(0.0015)     |
| Log petitions (rural non-pocketbook)            | 0.0004<br>(0.0010)                           | -0.0004<br>(0.0012)    | -0.0007<br>(0.0012)   | -0.0005<br>(0.0016)   | -0.0007<br>(0.0012)   | 0.0007<br>(0.0010)     |
| Log petitions (urban non-pocketbook)            | -0.0012<br>(0.0016)                          | -0.0019<br>(0.0018)    | -0.0022<br>(0.0018)   | -0.0027<br>(0.0029)   | -0.0022<br>(0.0018)   | -0.0014<br>(0.0016)    |
| Welfare topic share at $t$ (city GWR)           | -0.6814***<br>(0.0318)                       |                        |                       |                       |                       | -0.7337***<br>(0.0284) |
| Welfare topic share at $t - 1$ (city GWR)       |  | -0.1000***<br>(0.0290) |                       |                       |                       | 0.1243***<br>(0.0261)  |
| Welfare topic share at $t + 1$ (provincial GWR) |  |                        | -0.0248<br>(0.0539)   |                       |                       | -0.0058<br>(0.0436)    |
| Welfare topic share at $t$ (provincial GWR)     |  |                        |                       | -0.0495<br>(0.0518)   |                       | -0.0656<br>(0.0431)    |
| Welfare topic share at $t - 1$ (provincial GWR) |  |                        |                       |                       | 0.0280<br>(0.0470)    | -0.0439<br>(0.0394)    |
| Year fixed-effects                              | ✓  | ✓                      | ✓                     | ✓                     | ✓                     | ✓                      |
| Province-specific trends                        | ✓  | ✓                      | ✓                     | ✓                     | ✓                     | ✓                      |
| R <sup>2</sup>                                  | 0.38   | 0.05                   | 0.04                  | 0.04                  | 0.04                  | 0.40                   |
| Number of Cities                                | 297  | 295                    | 297                   | 297                   | 297                   | 295                    |
| Observations                                    | 1624   | 1553                   | 1624                  | 1624                  | 1624                  | 1553                   |

**Note:** This table shows regression results from including a variety of controls for the welfare topic share in provincial-level GWR or that in past city-level GWRs. The results indicate that our main coefficient of interest is highly robust to the inclusion of these potential confounders, suggesting that the effect of online participation is not endogenous to governments' prior policy orientations.

Robust standard errors clustered at city level are reported in parentheses.

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$  (two-tailed test)

## O Instrumental Variables Analysis

### O.1 Analysis Procedure

One possibility is that both policies and online participation may be driven by some unobserved time-varying events, such as a shift in societal preference for social welfare. We address this issue by looking for an instrument that is only correlated with online participation activities but uncorrelated with local demand for welfare. To do so, we exploit variations within petition topics. We construct an instrument for type RP petitions (rural, pocketbook) by calculating the average

number of type UN petitions (urban, non-pocketbook) from *all other* cities in the same province. Interviews with LLMB staff suggest that the public’s knowledge about the platform usually diffuses quickly within provincial boundaries, partly due to the relatively high frequency of communication and dense interpersonal networks within provinces. The average usage of the platform for non-welfare related issues should thus be a good proxy of the general popularity of the platform in a province. At the same time, the salience of rural pocketbook issues in a city should not, in theory, be directly related to the salience of urban, non-pocketbook issues in its neighbors because the two sets of issues concern different populations as well as different policy areas. While this exclusion restriction cannot be directly tested, we provide several tests in Section O.2 to address some potential challenges. For a city  $i$  in a province  $p$  with a total of  $N_p$  cities, the value of the instrument is calculated as:

$$IV_{ipt} = \log \left( \frac{\sum_{j \neq i} \text{Type UN Petitions}_{jpt}}{N_p - 1} \right).$$

Table A.13 presents the IV results for GWR topic shares. Column 1 shows that the first stage correlation between the two types of petitions is indeed very strong ( $F \simeq 200$ ), even though they cover very different issue areas. A one percent increase in type UN petitions from neighboring cities is associated with about 0.7 percent increase in RP petitions in the city of interest. The second stage results, shown in Column 2, are consistent with the OLS results: rural, pocketbook petitions appear to have a positive and statistically significant impact on changes in welfare topic shares. The magnitude of the IV estimate is slightly smaller than the baseline, indicating that concurrent offline activities might account for some, but not all the observed effects of online petitions. In the online appendix, we further show that the IV results are robust to using instruments constructed in various different ways (Table A.17 and Figure A.8).

Table A.13: Instrumental Variable Estimation

|  | $\Delta$ Welfare Topic at $t + 1$ (10-topic) |                       |
|--|--|-----------------------|
|  | (1)<br>First stage                           | (2)<br>Second stage   |
| IV: average petitions in neighboring cities (urban non-pocketbook) | 0.7394***<br>(0.0519)                        |                       |
| Log petitions (rural pocketbook)                                   |  | 0.0075***<br>(0.0023) |
| City and year fixed-effects  | ✓  | ✓                     |
| First stage F  | 203.27                                       |                       |
| Number of Cities   | 293  | 293                   |
| Observations   | 1620   | 1620                  |

**Note:** The table reports results from instrumental variables estimation. The dependent variables are incremental increase in the share of social welfare topic in government work reports. The specification is based on Column 3 of Table 2.

Robust standard errors clustered at city level are reported in parentheses.

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$  (two-tailed test)

Table A.14 reports the IV results on the *dibao* program, focusing on three metrics: individual coverage, family coverage, and total spending. Again, for all metrics, we see that RP petitions consistently have a positive and significant impact on the implementation of rural *dibao*: A one standard deviation increase in type RP petitions raises the program's individual coverage by about 1/7 of a standard deviation, family coverage by 1/4 of a standard deviation, and total spending by 1/5 of a standard deviation. These results are highly consistent with the OLS result reported in Section 5.5.

Table A.14: Change in Substantive Outcomes: Rural Minimum Living Standard Guarantee Scheme (*dibao*)

|                                  | Coverage (individual) | Coverage (family)     | Total spending        |
|----------------------------------|-----------------------|-----------------------|-----------------------|
|                                  | (1)                   | (2)                   | (3)                   |
|                                  | IV                    | IV                    | IV                    |
| Log petitions (rural pocketbook) | 0.0552***<br>(0.0193) | 0.0943***<br>(0.0219) | 0.0682***<br>(0.0226) |
| City and year fixed-effects      | ✓                     | ✓                     | ✓                     |
| First stage F                    | 224.43                | 224.43                | 224.55                |
| Number of Cities                 | 300                   | 300                   | 300                   |
| Observations                     | 1761                  | 1761                  | 1762                  |

**Note:** The table reports IV regression results using several implementation statistics of the rural *dibao* program as the dependent variable. The specification is based on Table A.13.

Robust standard errors clustered at city level are reported in parentheses.

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$  (two-tailed test)

## O.2 Evidence on Exclusion Restriction

A key condition for our instrumental variables analysis to be valid is the satisfaction of exclusion restriction—that is, the instrument does not affect our outcome of interest in ways other than through the independent variable. In this study, it means that the average volume of urban, non-pocketbook (UN) petitions from neighboring cities does not affect the a city’s welfare policy in ways other than affecting the general popularity of the LLMB in that city. There are two main threats to the exclusion restriction. First, the volume of UN petition may be correlated with certain provincial-level policies that affect all city governments’ policy priorities in the same province. It is plausible, for example, that a provincial administration that cares about lower-class citizens’ well-being would simultaneously encourage online petitions and ask its subordinate city governments to pay more attention to social welfare. One way to address this concern is to examine the relationship between UN petition and provincial governments’ policy orientations. If this channel of influence exists, we would expect the volume of UN petitions to be first associated with how provincial governments set their own welfare policies. Table A.15 reports regression results using several provincial-level welfare outcomes as the dependent variable. The first two columns use incremental changes in provincial GWR welfare topics, and the third through fifth columns report

results on fiscal spending patterns. Reassuringly, we do not find the volume of UN petitions to have significant correlation with any of the welfare outcomes. This suggests that provincial-level policy preference is unlikely to be a main confounder to our IV analysis.

Table A.15: UN Petition and Provincial Welfare Policy

|                                   | $\Delta$ Provincial welfare topic at $t + 1$ |                    | Provincial social expenditure at $t + 1$ |                     |                     |
|-----------------------------------|--|--------------------|--|---------------------|---------------------|
|                                   | (1)<br>10-topic                              | (2)<br>20-topic    | (3)<br>Welfare                           | (4)<br>Medical      | (5)<br>Agriculture  |
| Log UN petitions (provincial sum) | 0.0120<br>(0.0078)                           | 0.0102<br>(0.0077) | -0.0203<br>(0.0300)                      | -0.0160<br>(0.0298) | -0.0382<br>(0.0290) |
| Province and year fixed-effects   | ✓  | ✓                  | ✓  | ✓                   | ✓                   |
| R <sup>2</sup>                    | 0.13   | 0.10               | 0.99                                     | 0.99                | 0.99                |
| Observations                      | 144  | 144                | 144                                      | 144                 | 144                 |

**Note:** The table reports the association between the total (log) number of urban non-pocket petition in a province and provincial welfare policies. Provincial gdp and population are included as controls. Robust standard errors clustered at province level are reported in parentheses.

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$  (two-tailed test)

The second potential threat to our exclusion restriction is that changes in UN petitions in a city are directly related to welfare policy changes in that city, which would then affect policies in other cities through regional diffusion. For this mechanism to hold, we would expect to first observe a positive correlation between a city's UN petitions and its own welfare policies.<sup>32</sup> To test this, we regress a number of city-level welfare outcomes on a city's own level of UN petitions. The outcomes include GWR welfare topics, fiscal expenditure, and coverage of the Minimum Living Standard Guarantee Scheme (also known as *dibao*, see Section 5.5 for more information). As shown in Table A.16, the correlation between a city's UN petitions and its own welfare outcomes are actually rather weak. This provides evidence against the diffusion hypothesis.

<sup>32</sup>It is worth noting that correlation does not have a transitive property (see Langford, Schwertman, and Owens (2001) for a formal proof). In other words, even the volume of UN petitions is likely to be positively correlated with pocket petitions and pocketbook petitions are positively correlated with welfare policies, this does not imply that UN petition will necessarily be positively correlated with welfare policies.

Table A.16: UN Petitions and A City's Own Welfare Outcomes

|                                      | $\Delta$ Welfare Topic |                    | Fiscal Expenditure |                     |                    | <i>Dibao</i> Coverage |                     |
|--------------------------------------|------------------------|--------------------|--------------------|---------------------|--------------------|-----------------------|---------------------|
|                                      | (1)<br>10-topic        | (2)<br>20-topic    | (3)<br>welfare     | (4)<br>medical      | (5)<br>agiculture  | (6)<br>urban          | (7)<br>rural        |
| Log petitions (urban non-pocketbook) | 0.0022<br>(0.0020)     | 0.0024<br>(0.0018) | 0.0007<br>(0.0123) | -0.0110<br>(0.0112) | 0.0114<br>(0.0142) | -0.0055<br>(0.0110)   | -0.0144<br>(0.0103) |
| Province and year fixed-effects      | ✓                      | ✓                  | ✓                  | ✓                   | ✓                  | ✓                     | ✓                   |
| R <sup>2</sup>                       | 0.03                   | 0.03               | 0.83               | 0.86                | 0.84               | 0.41                  | 0.37                |
| Observations                         | 1624                   | 1624               | 1778               | 1777                | 1778               | 1762                  | 1778                |

**Note:** The table reports the association between urban non-pocketbook petitions and a series of welfare-related outcomes at the city level. The specification is based on Column 3 of Table 2. Robust standard errors clustered at city level are reported in parentheses.

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$  (two-tailed test)

### O.3 Robustness Checks and Placebo Tests

We use topic classifications from alternative LDA models to replicate the IV analysis. To do so, we estimate a 20-topic and a 40-topic LDA models, group the topics into the four categories (RP, UP, RN, UN) as before, and then perform the same estimation, instrumenting RP petitions with the volume of UN petitions from neighboring cities. As shown in the top and middle panels of Table A.17, the main IV results remain largely unchanged. In the bottom panel of Table A.17, we use the volume of UN petitions in the provincial capital as an alternative instrument, and find largely similar results.

Table A.17: Alternative Instruments

| Instrument: <i>Urban, Non-Pocketbook Petitions (UN) from Other Cities in the Same Province (20-Topic Model)</i> |                               |   |                                       |                                    |
|---|-------------------------------|---|---------------------------------------|------------------------------------|
|   | (1)<br>$\Delta$ Welfare topic | (2)<br><i>Dibao</i> Coverage (individual) | (3)<br><i>Dibao</i> Coverage (family) | (4)<br>Total <i>dibao</i> spending |
| Log petition (rural pocketbook, 20-topic model)   | 0.0074***<br>(0.0022)         | 0.0529***<br>(0.0187)                     | 0.0950***<br>(0.0214)                 | 0.0646***<br>(0.0216)              |
| City and year fixed-effects   | ✓                             | ✓   | ✓                                     | ✓                                  |
| First stage F   | 231.13                        | 245.68                                    | 245.68                                | 245.69                             |
| Number of Cities  | 293                           | 300                                       | 300                                   | 300                                |
| Observations  | 1620                          | 1761                                      | 1761                                  | 1762                               |
| Instrument: <i>Urban, Non-Pocketbook Petitions (UN) from Other Cities in the Same Province (40-Topic Model)</i> |                               |   |                                       |                                    |
|   | (1)<br>$\Delta$ Welfare topic | (2)<br><i>Dibao</i> Coverage (individual) | (3)<br><i>Dibao</i> Coverage (family) | (4)<br>Total <i>dibao</i> spending |
| Log petition (rural pocketbook, 40-topic model)   | 0.0075***<br>(0.0022)         | 0.0525***<br>(0.0190)                     | 0.0911***<br>(0.0214)                 | 0.0642***<br>(0.0221)              |
| City and year fixed-effects   | ✓                             | ✓   | ✓                                     | ✓                                  |
| First stage F   | 202.15                        | 222.66                                    | 222.66                                | 222.82                             |
| Number of Cities  | 293                           | 300                                       | 300                                   | 300                                |
| Observations  | 1620                          | 1761                                      | 1761                                  | 1762                               |
| Instrument: <i>Urban, Non-pocketbook Petitions (UN) in Provincial Capital</i>                                   |                               |   |                                       |                                    |
|   | (1)<br>$\Delta$ Welfare topic | (2)<br><i>Dibao</i> Coverage (individual) | (3)<br><i>Dibao</i> Coverage (family) | (4)<br>Total <i>dibao</i> spending |
| Log petitions (rural pocketbook)  | 0.0080***<br>(0.0027)         | 0.0239<br>(0.0245)                        | 0.0443*<br>(0.0248)                   | 0.0624**<br>(0.0262)               |
| City and year fixed-effects   | ✓                             | ✓   | ✓                                     | ✓                                  |
| First stage F   | 138.53                        | 165.61                                    | 165.61                                | 165.73                             |
| Number of Cities  | 293                           | 300                                       | 300                                   | 300                                |
| Observations  | 1620                          | 1761                                      | 1761                                  | 1762                               |

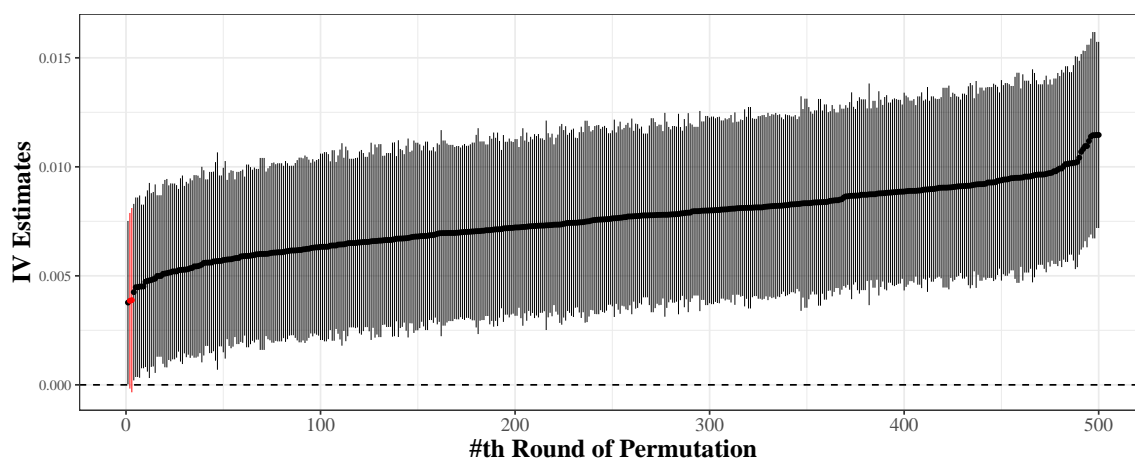
**Note:** The table reports replication of the main topic and spending results using measures generated from a 40-topic LDA model. The specification is based on Column 3 of Table 2. Robust standard errors clustered at city level are reported in parentheses.

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$  (two-tailed test)

To further address the issue that our results are driven by unobserved correlations between rural, income-related petitions and certain specific types of petitions that we use as instruments,

we also run a permutation tests where we *randomly* select petition volumes for three non-rural, non-pocketbook topics to construct the instrument and re-estimate the IV regression. We repeat this exercise for 500 times and visualize the results in Figure A.8. It turns out that the estimated coefficients are positive in all 500 iterations and miss the 10% threshold of significance in only 2 iterations. This permutation exercise strengthens our confidence that our IV results are not driven by the relationship between rural pocketbook petitions and any specific urban non-pocketbook topic.

Figure A.8: Permutation of Topics Used for Instrument



*Note:* This figure visually presents the results from 500 permutations of the instrument. Only 2 out of 500 permutations turn out to be statistically insignificant at 10% level (highlighted in red)

Moreover, one may be worried that the diffusion of LLMB might be correlated with provincial-level shocks to welfare policies. To check whether this is the case, we conduct a placebo test using the implementation of the *dibao* program in urban areas as the alternative outcome. The results are shown in Table A.18. As expected, rural pocketbook petitions have little effect on the distribution of urban *dibao*. This gives us confidence that our finding is not driven by unobserved correlations between platform diffusion and governments' preferences for welfare policies.



Table A.18: Placebo using Urban Dibao as Outcome

|                                  | No.<br>individuals<br>covered (log) | Total<br>spending<br>(log) | Per capita<br>spending<br>(log) |
|----------------------------------|-------------------------------------|----------------------------|---------------------------------|
|                                  | (1)                                 | (2)                        | (3)                             |
|                                  | IV                                  | IV                         | IV                              |
| Log petitions (rural pocketbook) | 0.0012<br>(0.0182)                  | -0.0308<br>(0.0197)        | -0.0163<br>(0.0110)             |
| City and year fixed-effects      | ✓                                   | ✓                          | ✓                               |
| First stage F                    | 229.54                              | 229.54                     | 229.54                          |
| Number of Cities                 | 302                                 | 302                        | 302                             |
| Observations                     | 1777                                | 1777                       | 1777                            |

**Note:** The table reports a placebo test using urban *dibao* spending as the outcome. The result suggests that rural pocketbook petitions have no effect on the implementation of urban *dibao*. The specification is based on Column 3 of Table 2. Robust standard errors clustered at city level are reported in parentheses.

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$  (two-tailed test)

## P Extensions

### P.1 Effects of Online Petitions on Spending Patterns

Table A.19: Effects of Online Petitions on Fiscal Expenditures

|                             | Growth in Expenditure |                     |                    |                     |
|-----------------------------|-----------------------|---------------------|--------------------|---------------------|
|                             | (1)                   | (2)                 | (3)                | (4)                 |
|                             | social                | medical             | education          | science             |
| Log petitions               | 0.0570<br>(0.0452)    | -0.2215<br>(0.2124) | 0.0326<br>(0.0227) | -0.0012<br>(0.0319) |
| Year and city fixed-effects | ✓                     | ✓                   | ✓                  | ✓                   |
| Economic controls           | ✓                     | ✓                   | ✓                  | ✓                   |
| Leadership controls         | ✓                     | ✓                   | ✓                  | ✓                   |
| R <sup>2</sup>              | 0.16                  | 0.02                | 0.13               | 0.04                |
| Number of Cities            | 303                   | 303                 | 303                | 280                 |
| Observations                | 1773                  | 1777                | 1762               | 1637                |

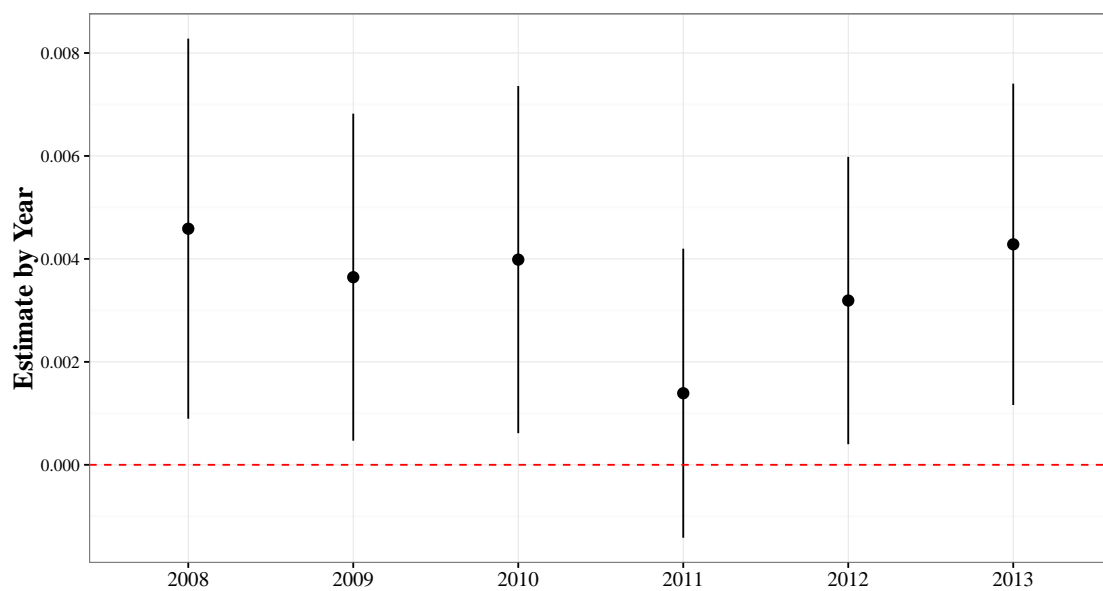
**Note:** The table reports the estimated effects of online petition on government expenditure in several areas. The specification is based on Column 3 of Table 2. Robust standard errors clustered at city level are reported in parentheses.

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$  (two-tailed test)

## P.2 Effects of Participation over Time

We also investigate the possibility that our results were driven by rare events that happened in a few special years. To do so, we estimate a model where the participation variable is interacted with a set of year dummies. It is clear that the effect is in fact quite stable over time. All estimates are statistically significant at 10% level or higher with the exception of 2011. This pattern lends us further confidence that the result is not driven by the choice of sample period.

Figure A.9: Effect of Online Participation by Year



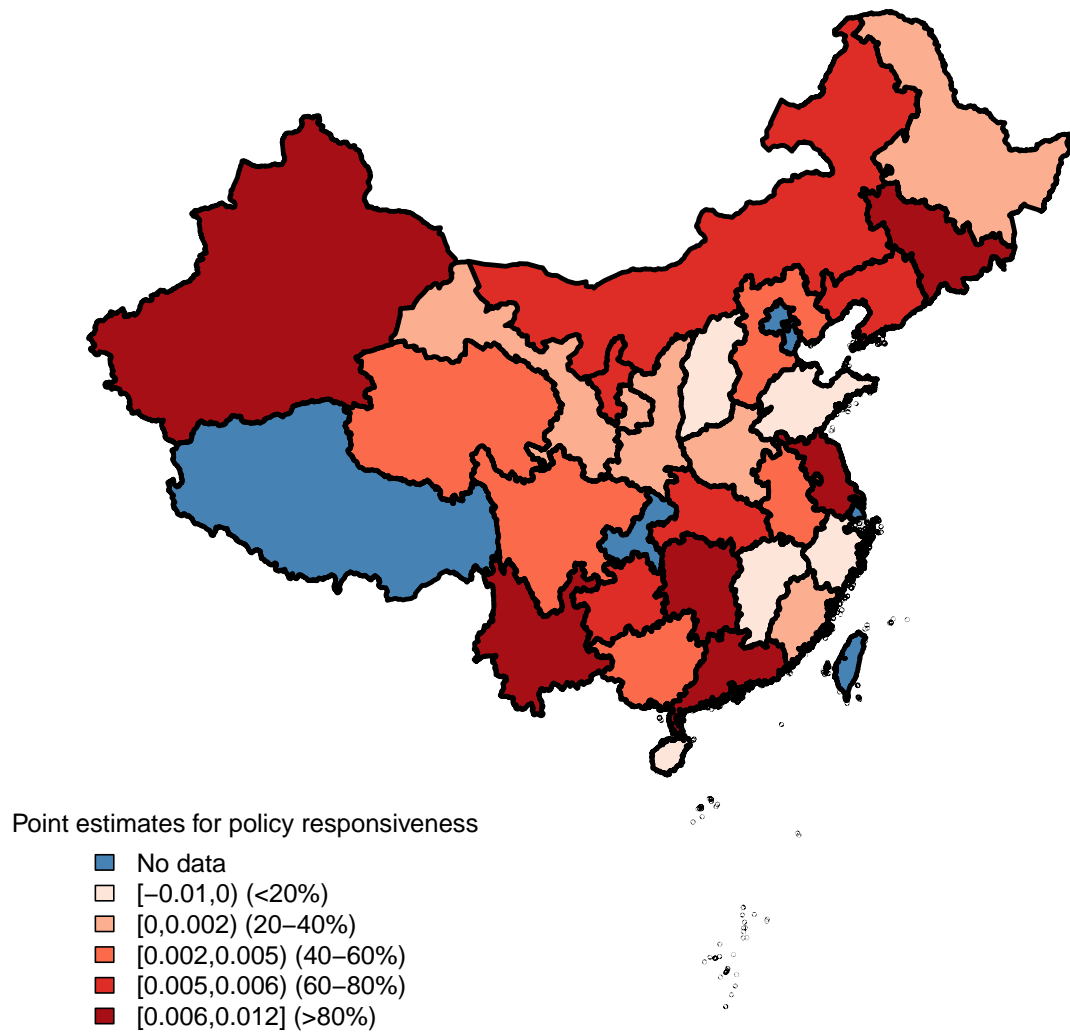
### P.3 Geographic Variation in Responsiveness

We also estimate province-specific policy responsiveness using the following equation:

$$\Delta \text{Welfare Topic Share}_{it+1} = \alpha \text{Log Petitions}_{it} + \delta^p \text{Log Petitions}_{it} \times \text{Province}_p + \mathbf{X}_{it}\boldsymbol{\beta} + \phi_i + \tau_t + \epsilon_{it}.$$

The level of policy responsiveness for province  $p$  is thus  $\alpha + \delta^p$ . We plot this quantity in the following map.

Figure A.10: The Geography of Policy Responsiveness



## Q Fieldwork Evidence on the Petition-Policy Link

In this section, we draw on materials from our fieldwork to flesh out the process by which on-line petitions affect government policy making. Between 2013 and 2017, we conducted several dozens of interviews with central and local government officials and staff at the People Online (the company that manages LLMB) to learn about how Chinese government handles online petitions. According to our interviews, most local governments today have set up specialized agencies dedicated to monitoring social media sites and online petition platforms. These agencies, usually under names such as Public Opinion Office (輿情辦公室) and Social Condition and Mass Opinion Office (社情民意辦公室), are responsible for collecting information about online public opinion from popular websites and reporting it to key local decision makers (e.g., party secretary, mayor, or members of the local party standing committee). The reporting takes several forms: The first are daily, weekly, or monthly briefings that summarize trending issues in online discussion. These reports usually cover a wide range of policy or non-policy topics and are intended to keep local leaders updated about changes in the focus of public attention. A second type of reports are in-depth investigation on public reactions to specific accidents or events, such as scandals, collective protests, or safety accidents. This type of reports usually combines online public opinion information with on-ground research to produce actionable plans for local authority to manage and direct public sentiment. Finally, a third type of reports are policy recommendations based on public opinion in a specific policy domain (e.g., education, health care, or social welfare). When drafting guiding policy documents, such as the Government Work Reports, local governments will consult a wide range of agencies and actors, and input from online public opinion has become increasingly important in recent years as the influence of online platforms grow.<sup>33</sup>

Our fieldwork also suggests that the specific ways by which local governments handle online petitions vary greatly across localities. Usually the handling of online petitions is led by one the following agencies: the general office of party committee (黨委辦公室), the government general

---

<sup>33</sup>The central government, for example, initiated a campaign entitled “I spoke to the Premier” to encourage netizen input to the drafting of the Central Government Work Report in 2014. Many local governments have also made their own initiatives. See <https://goo.gl/d1oDQP>; <https://goo.gl/qwS8SG>; <https://goo.gl/nyFHAC>.

office (政府办公室), the cyber administration office (网信办), the publicity bureau (宣传部), the supervision department (监察局), or the letters and visits bureau (信访局). The most popular choice is to manage online petitions through the general office of the local party committee; this practice are adopted by eight provinces, including Zhejiang, Anhui, and Gansu. Another popular model uses the government general office as the leading management agency; Tianjin, Shanghai, and Liaoning utilize this mode. In some other provinces, such as Beijing and Shandong, the cyber administration office is in charge of monitoring petitions and issuing replies. In Hunan, online petitions are managed by the supervision department. In Fujian, the letters and visits bureau is authorized to manage the online petitions. Shanxi, moreover, created a new omnibus agency called public opinion office to carry out all tasks related to online public opinion management.

## References

- Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4), 77–84.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3, 993–1022.
- Langford, E., Schwartzman, N., & Owens, M. (2001). Is the property of being positively correlated transitive? *The American Statistician*, 55(4), 322–325.
- McCallum, A. K. (2002). *Mallet: A machine learning for language toolkit*. <http://mallet.cs.umass.edu>.
- Mimno, D., Wallach, H. M., Talley, E., Leenders, M., & McCallum, A. (2011). Optimizing semantic coherence in topic models. In *Proceedings of the conference on empirical methods in natural language processing* (pp. 262–272). EMNLP '11. Edinburgh, United Kingdom: Association for Computational Linguistics.
- Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica*, 49(6), 1417–1426.
- qinwf. (2016). *Jiebar: Chinese text segmentation with r*. <https://github.com/qinwf/jiebaR>.