

Automated Evaluation of Deputy Social Responsiveness in China's People's Congress

Xingjian YAN

Abstract: Increasing deputies' social responsiveness is a principle task for China's legislatures as it is remarkably weakened by the lack of competitive elections that is *de facto* an appointment and selection process. This article applies "word score" computer-aided method to the study of the 'deputy proposals' submitted by the deputies from a selected city People's Congress (PC). The data illustrates that more than a half of the proposals have strong linkage to public concern and local demands (e.g., to increase education fund). This proportion is higher than those of provincial and national PCs. However, data also suggests a lack of direct deputy-voter contact, as a great number of the proposals are based on the deputies' personal experience or professional expertise, rather than the voters' demand. This research is also a preliminary attempt to applying the automated content analysis to the PC proposals that, given Chinese language's high 'drifting' nature (a word's meaning changes dramatically within different context), were previous not subject to any automated method to analyse. In order to test and refine the method, the texts subject to the automatic coding are also manually coded. Comparing the two outcomes, the paper suggests three ways to enhance accuracy. This paper concludes by suggesting that the refined method can be applied to a large quantity of deputy proposals (even in national scale), by which a database can be established and facilitate further studies.

Key words: People's Congress; Deputy Responsiveness; Content Analysis, Word Score

1. Research Background

Legislatures connect the public and the political system. In the case of China, the representation function of the People's Congress (PC, consists of five layers: national, provincial, city, county, and township) relies on a series of connection that allow the information to exchange in both vertical and horizontal direction. There are two most important types of connection, namely, the PC-deputy connection and deputy-voter connection. This paper focuses on the second one that aims to collect and report public demands and grievances to the political system.

This connection depends on deputies' active responsiveness to the public demands; however, a lack of responsiveness is always seen throughout the PC history (O'Brien 1994, Truex 2014). The principle reason roots in the voting

system. The extremely limited competitiveness, the ban of electoral campaign, and the non-general election cut off the accountability tie between the deputy and voters (Cai, 2003).¹

Deputies' social responsiveness is also damaged by a unique philosophy of representation that can partially be compared to 'descriptive representation' (Pitkin 1967). This type of representation stresses that a deputy bearing a certain character is the prerequisite that he or she will fight for the interest of the people with the similar character (e.g., women deputy will guard the interest of women in the legislature). In China, this philosophy is further developed into an attempt that the deputies' social characters, especially their work unit and expertise, should 'mirror' the society. This led to both institutional and behavioural sequences. The institutional outcome is a series of institutional design that emphasises deputy-work unit connection at the price of reducing deputy-voter connection (e.g., deputies are part time; very short PC session). The behavioural outcome is the deputies' preference to introduce bills or proposals that has relationship to their work or profession, not their constituents.

Recent research applies two approaches to explore the PC deputies' connection to the public. The 'empirical' approach is featured by the quantitative methods that include interviews with deputies and mass scale voter surveys. It is a case oriented approach that features describing actual behavioural changes of the deputies. Recent studies taking this approach reveal that at present a small number of the PC deputies start to pay greater attention to the demands and grievances from the public (He & Wang 2012, Manion 2014a, Manion 2014b). However, further examinations show that such a change is resulted from a rising number of deputies with changed social background (younger age, higher income, and experts from some professions), rather than any significant institutional change. For example, a greater number of deputies' higher education background brings an indigenous sense of accountability to some extent (Huang & Chen 2015). The 'text' approach focuses on deputies' formal bills (Yi An) or proposals (Jianyi). The artificial reading and coding of text is used as an alternative in case that (1) the empirical approach is too case sensitive to draw any general conclusion, (2) interview or survey data normally is not available or difficult to obtain (Sang & Qiu 2010, Xu 2013). However, the major handicap of this approach is that the number of cases a study may involve is rather limited by time.

Both approaches therefore limit the number of cases. It is apparent that there should be a method that can be applied to automatically evaluating a large quantity of data within a reasonable amount of time and cost. The proposed method might help legislative scholars create a provincial level or even national

¹ Among the five levels of the PC, only the last two are general elected. Deputies of other three levels are elected by the deputies of lower PC.

level database, where the deputies' responsiveness of every city or county is evaluated and collected. This method should also overcome the limitation of interview and questionnaire methods that are poor in mining past evidence (say, how was the deputies' responsiveness 5 years ago?). Panel data might further contribute to advancing the understanding of the progress and limitations of the deputies' responsiveness to public concerns, and shed light to future institutional developments.

An automated analysis of deputies' formal bills and proposals matches the purposes discussed above. Formal bills and proposals are structured texts that are suitable for automated and quantitative based analysis. The rationale for this process is automatically coding the deputies' formal bills and proposals in order to reveal to what extent they focuses on local demands or grievances. Such an extent shows how active a deputy responses to the voter's concern, which is an ideal measurement to the deputies' responsiveness.

2. 'Word score' method

Methods in analysing text can be categorised into two types, namely the qualitative 'discourse analyses and the quantitative 'content analyses'. The quantitative nature of the second type enables automated analysis, as it treats words from text as pure codes, instead of trying to 'decipher' the meaning through understanding its cultural background, power structure, or any qualitative implication. Computer assisted Automated Content Analysis (ACA) has been widely used to the analysis of a large quantity of structured text in recent years. Despite the computer assisted 'word frequency' method, perhaps the easiest content analysis method, is still widely applied in social science research, more sophisticated methods have been developing since the 1970s, some appear to be effective in analysing various parliamentary documents, including legislative bills, Hansards of parliamentary debate, and MP - voter correspondence (Quinn *et al.* 2010, Sieberer *et al.* 2016).

ACA methods consist of two pathways of text processing, namely the 'known categories' and 'unknown categories'. The former pathway consists of various dictionary methods that apply pre-defined artificial coding dictionary to score texts. The latter consists of various supervised methods. 'Supervised' is named after a 'manual training' process: a certain number of texts are selected, manually assessed, and then used to train computer. In the next step, a computer uses this 'ruler' to automatically analysing a large quantity of texts (Grimmer & Stewart 2013, Simons & Xenos 2004).

Compared with Latin system, Chinese language adds difficulty to ACA, as researchers have to combine individual characters to shape words in order to get meaning. For example, as there is no space to 'gap' the words, when a computer

reads a word that is composed of three characters (say, abc) without enough context to understand, it either reads it as 'a bc' or 'ab c', and gets completely different meaning. Therefore, manual training (supervised) pathway is more suitable for Chinese language, as the manual training process can reduce the risk of misunderstanding by computer. This study applies 'word score' method of this pathway.

A typical 'word score' analysis includes three steps. First, all of the texts to be evaluated are divided into two groups, namely the 'training texts' and 'evaluating texts'. Training texts are read and scored manually. Second, a series of words are selected from the overall text. Computer aided calculation counts the number of appearance (word frequency) of each selected word in each of the training text. For each of the word A, B, C..., the formula (1) shown below is applied to link the word and the emotion or function of the text. The basic rationale for this process is that if a word has a very high frequency of appearance in a text, and that text has been confirmed to express a certain emotion or function in the manual evaluating process, then the high frequency of appearance of this word in any text indicates that this text may express the same emotion or function. Table 1 provides an example of scoring the three words.

Table 1: scoring of the words

	Word frequency*			Word score			
	Text 1	Text 2	Text 3	Text 1	Text 2	Text 3	Final score
Word A	2 (1.00)	0 (0.00)	0 (0.00)	-2	0	0	-2
Word B	4 (0.67)	2 (0.33)	0 (0.00)	-1.33	0.17	0	-1.16
Word C	3 (0.75)	0 (0.00)	1 (0.25)	-1.5	0	0.25	-1.25
Text score	-2	0.5	1				

*) showing each the word's frequency and percentage distribution in 3 texts

Source: Author

The third step is to use the words that have been scored to measure the 'evaluating texts' (formula 2). Once a research has manually scored every selected word, a computer can use them as a ruler (so called 'ruler words') to analyse a huge amount of text within a very short period of time, given that the texts submitted for the automated analysis belongs to the same type of the texts that are used as 'training text'. The basic rationale of this process is to evaluate the texts' function or emotion by considering the frequency of the ruler words. For example, suppose that a ruler word A, being confirmed to have strong linkage to a text's type 1 emotion or function, appears 10 times in a text, while word B, being confirmed to have strong linkage to a text's type 2 emotion or function, appears only 1 time. It could be inferred that the text might be sorted

into type 1 emotion or function.

$$As = \sum_{i=1}^r \left(\frac{ANr}{\sum_{i=1}^r AN} \right) \times Pr \quad (1)$$

$$St = \sum_{i=1}^A (ANt / \sum_{i=1}^A Nt) \times As \quad (2)$$

ANr the word frequency of the ruler word 'A' in training text 'r'.

$\sum_{i=1}^r AN$ the word frequency of the ruler word 'A' in all training texts.

Pr the score of text 'r' from the manual coding process

ANt the word frequency of the ruler word 'A' in the evaluating text 't'

$\sum_{i=1}^A Nt$ the word frequency of all ruler words in the evaluating text 't'

3. Research design

Based on the procedurals shown above, the data is processed by the following steps:

(1) Case: the text data for this research is drawn from the deputies' formal bills and proposals of N City People's Congress. N city locates in the southeast of China, is a provincial capital. About 180-250 proposals and 3-10 formal bills submitted in each annual plenary session conference (lasts about 6 days).² This research only analyses the proposals, as the number of formal bills is too few. Deputy responsiveness is crucial for the PC of lower levels (city, county, and township). Compared with higher levels, the PC of the city and the lower levels have strong linkage to the public and individual voters, and should be an important channel for the public to express their demands and grievances.

The first annual session of the 14th PC (2011) is used in this study as the training text; Proposals submitted to the fourth annual session of the 13th PC and the second session of the 14th PC are used as 'evaluation texts'. Each proposal is processed' as one case (also referred to one 'text'). The number of cases from the three sessions of the conference are 181, 226 and 173 respectively.

(2) Ruler words: The 'word score' method estimates text functions or emotion through words. Therefore, the more the ruler words being used to estimate, the more accurate the estimation is. In idea situation, it requires all words from

² Each People's Congress lasts five years, holding approximately five plenary session conferences. Each conference lasts approximately 6 days, during which deputies submit their Yian and Proposals. Corresponding governmental departments and other organisations are required to respond.

training texts that might be relevant to the text function to be included (Laver *et al* 2003). However, the deputy proposals this research uses have three common features: (1) short length (normally around one to two hundred Chinese words); (2) function (all aim to express opinions, suggestions, demands to the corresponding PC), and (3) the use of language (texts use formal vocabulary instead of any literary vocabulary, at some point resulting in a fixed linkage between vocabulary and text function).

This research therefore manually chooses words for scoring from three broad categories that echo the deputies' three possible roles, which are 'responding to public concern', 'professional consultant', and 'overall and overarching issues'. The author and two postgraduate research assistants have read, discussed, and selected the training texts. This generates a raw wordlist containing 30 words. We subsequently counted their frequency in each training text and evaluation text. This process removes some of the words that either appear too uncommon (the total number of appearance is less than 50) or have approximately equal frequency in each text (this type of words are weak in pointing out the function or the position of the text, as their word scores always approach the middle session of the coding spectrum)

Table 2: list of ruler words

	Responding to public concern	Professional consultant	Overall and overarching issues
remain	Resident, Citizen, Masses, Region	Occupation (Hangye), Industry, Department, Work unit	Development, Promotion, the Whole City, Problem, Solution
remove	Local (Dangdi),* Local (Bendi), Regional, People, area	Occupation (Zhiye), Profession, Field	Reform, Whole, Advance, Innovation

Source: Author

* English translation may have difficulty reflecting the nuance between two Chinese words. If two Chinese words share one English translation, Pingying (Phonic transcription of Chinese words) is added to distinguish.

(3) Creating coding spectrum and the corresponding coding scheme, by which the training text is manually scored. Scoring of the training text is the key step for 'word score' analysis. The accuracy of the scoring procedure is secured by introducing double-checking strategy: The author and two postgraduate research assistants scored the text independently from each other. Based on the result, an average score is calculated. The author also re-scored the texts with significant deviated scores (larger than 1) given by different examiners. The

simple reliability ratio is between 70-80%.³

The coding scheme consists of five marks: -2, -1, 0, 1, 2. The basic logic is that texts aiming to highlight and report constituency claims, grievances, or problems are given positive scores, whereas a negative score is given. The simplified coding scheme is shown below (table 3). The scheme was amended several times during the pre-coding section to avoid any overlapping of the criteria.

It is worth noting that the positive or negative score in the scheme does not suggest a deputy proposal highlighting constituency claims is more valuable than that focusing on a grand or overarching issue. Deputies using their professional skills or knowledge to provide their ideas to an overarching issue should not be simply equated to a weak deputy responsibility. However, given the fact that a lack of deputy-constituency connection and the deputies' responsiveness to the constituency claims is the major challenge to the PC, encouraging deputies to pay greater attention to constituency issues is therefore an expected orientation of development. It is particularly so for the local level PCs deputies. In this case, the city level PCs deputies' connecting to every individual constituent is a crucial bridge through which the executive could hear the voice and claims of the public, by which the legislatures' safety valve function is fulfilled.

Table 3: Coding Scheme

Question: to what extent do you agree that the deputy proposal aims to express constituency claims or grievances?		
Response	Score	Criteria (simplified)
Totally Agree	2	Proposal aims to highlight specific constituency claims, grievances, or problems that is directly expressed by the constituents
Partially Agree	1	Proposal aims to highlight a specific issue in the constituency; however, the issue is neither raised by the voters, nor highlighted by the deputy-voter connection, but being noticed by the deputies through their own experience or professional expertise.
Yes, but not the claim by constituency	0	Proposal aims to highlight a specific issue of public concern, but the issue has strong linkage to the deputies' profession or work unit.
Partially Disagree	-1	Proposal aims not to highlight a specific constituency issue, but to introduce a plan for overarching development (i.e. construction of a railway), or the proposal is of 'government work report' style that aims to advertising the social development of the past year or a specific period of time.
Strong Disagree	-2	Proposals having no feasibility, as it (1) introduces a massive plan or project that surpasses the authority or capability of city level

³ Reliability ratio is calculated by dividing the number of texts receiving consistent score (diversity of scores a text receives is lower than 1) from different examiners by the total number of texts.

		government; (2) is too empty and cannot be forwarded to a government department to process; (3) is a 'ridiculous' proposal (see specific discussion below)
--	--	--

Source: author

(4) The score for each selected word is calculated using formula 1. Based on the scores of the 15 words, formula 2 is applied to scoring the 'evaluation texts'.

(5) Data test and analysis. Content analysis explores the linkage between words and the function of the text. Such a linkage is largely built on statistical significance, rather than logical connection, given the fact that the words in Chinese language are too 'versatile'. The meaning of a word can change dramatically in different contexts. The accuracy of the word score method should be tested. In so doing, we manually scored the 'evaluation texts' that had been evaluated by the automated method. We subsequently compared the outcome of the two methods to discuss any improvements on methodology.

3. Data analysis and methodological consideration

The scores for the 15 words are given in table 4, including the word frequency in all 181 training texts, word score, minimum score, maximum score, and the stand deviation that evaluates the degree of dispersion for each word.

Table 4: Scores of the selected words

Word	Frequency	Score	Min.	Max.	Sdv
Resident	81	1.0154	-0.0432	0.0222	0.0271
Citizen	101	0.5743	-0.0792	0.2574	0.0239
Masses	76	0.3816	-0.0132	0.0789	0.0216
Region	65	-0.3038	-0.1846	0.0923	0.0247
Occupation	59	-0.8686	-0.4746	0.0678	0.0461
Industry	148	-1.4561	-0.6351	0.0405	0.0525
Department	213	-0.2230	-0.1221	0.0798	0.0173
Work unit	70	-0.0214	-0.2000	0.1000	0.0230
Development	456	-0.4709	-0.0965	0.0439	0.0169
Promotion	70	-0.2000	-0.1714	0.1714	0.0222
Whole city	66	-0.1629	-0.1818	0.0909	0.0243
Problem	165	0.1152	-0.0788	0.0848	0.0178
Solution	107	0.0888	-0.0935	0.1495	0.0233
Economy	253	-0.5484	-0.1818	0.0346	0.0209
Environment	143	-0.1031	-0.1678	0.0629	0.0207

Source: author

The score for each ruler word met our expectation. For example, three of the ruler words ‘citizen’, ‘masses’, and ‘resident’ are expected to be able to illustrates the text’s function of expressing public demands. In this study, all three words received positive score.

The scores of the 15 words are applied to the automated evaluation of the deputy proposals from the second plenary session of the 14th PC, and the 4th plenary session of the 13th PC. All of the proposals are also manually evaluated so as to test the reliability of the word score method.

Table 5 compares the outcome of the two methods. Basically, the tiny difference between the two outcomes suggests that the word score method is reliable in evaluating the function or position of the deputy proposals. In this research, the difference of two methods in evaluating the 2nd session of the 14th PC is only 1.7%. Some of the 13th PC deputies were replaced by new deputies in the 14th PC, the different behaviour of language use (use of vocabulary in particular) reduced the accuracy of the automated method that the ‘ruler’ is on the basis of the 14th PC. The difference of the 4th session of the 13th PC is therefore higher.

Table 5: Comparing the outcome of two methods

Source of Text	Proposals with positive score (total number and percentage)		Difference
	automated	manual	
14 th PC 2 nd session	100 (57.8%)	97 (56.1%)	1.7%
13 th PC 4 th session	119 (52.7%)	107 (47.4%)	5.3%

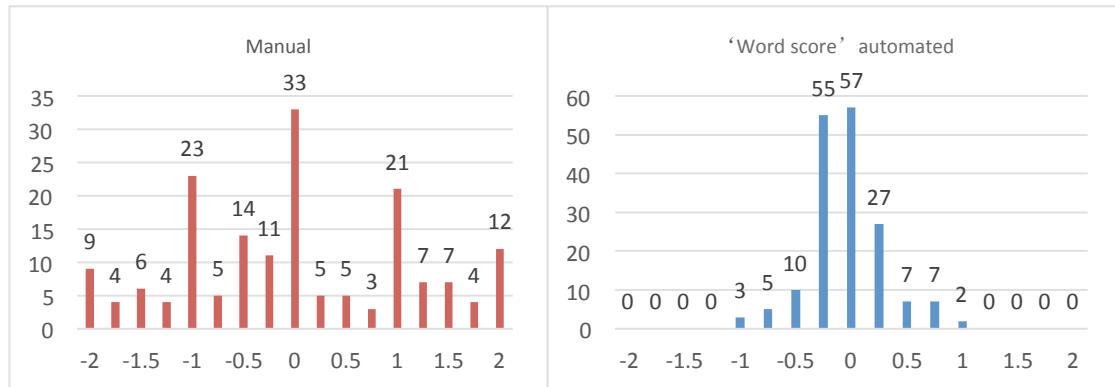
Source: author

However, word score method is strong in evaluating the entire sample, but weak in evaluating the individual case. Taking the proposals from the 2nd session of the 14th PC as an example. Some cases receive very different score by two methods, thus raise all cases’ athematic mean of the diversity between two methods to 0.74.⁴ The reason for the difference is understandable by comparing the frequency distribution of the cases (Chart 1): word score method results in a unimodal normal distribution and peaks at the ‘0-0.25’ interval; the largest number of cases (almost 40 percent) appear in the ‘-0.25-0.5’ interval. The manual evaluation results in a multimodal normal distribution; although the peak is also at “0-0.25” interval; the number of cases in this interval is clearly less than that from the automated method. Moreover, not a single case, in the automated method, appears in the intervals that are lower than -1 or higher than 1.25. In contrast, the outcome of manual coding put 12 cases in the highest

⁴ $(\sum_{i=1}^n [\text{score of automated evaluation} - \text{score of manual evaluation}]) / n$, where ‘n’ consists of all cases from the 14th PC 2nd session and the 13th PC 4th Session that are both automatically and manually scored.

interval.

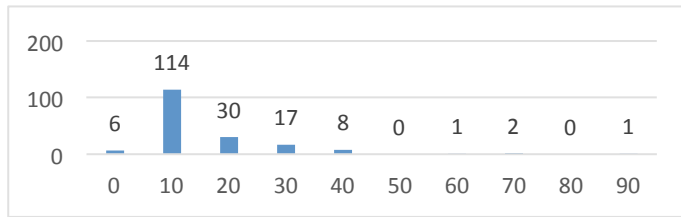
Chart 1: Frequency distribution of the texts' scores by two methods



The principle reason for the diversity given above is the mechanism of the two methods. The fact that many cases are allocated to the two edges of the frequency distribution spectrum by the manual method illustrates a key weakness of the word score automated method: words are judged by quantitative word frequency, rather than qualitative meaning in the text. For example, in manual scoring, a single word from the sentence “local people expressed their strong antipathy towards this problem” results in a deputy proposal receiving a score of 2; however in the automated method, although the word ‘antipathy’ and ‘local people’ may lead the text to receive a positive score, the text cannot be pushed to the pole of the spectrum unless such words repeat several times.

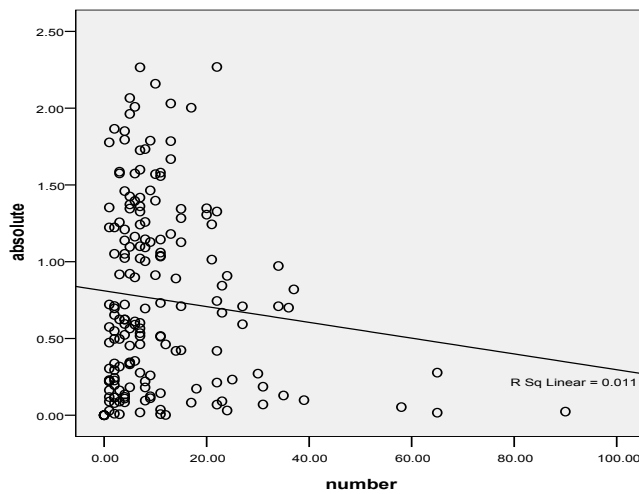
In addition, the process of choosing and scoring the ruler words also magnified the diversity. First, the number of ruler-word is too little. As previously mentioned, word score method measures the quantity of words. Too few words reduces the chance for a proposal to receive very high or low score. According to chart 2, the vast majority of the text contains only 10-20 measurable words (words that matches the ruler-word list and can be used to score the text). Chart 3 explores the correlation between the number of each text’s measurable words and the difference between two methods. The overall correlation is not significant ($R^2=0.011$); however, once the number of measurable word in one text exceeds 25, the correlation becomes more significant ($R^2=0.312$, and the difference by two methods is always lower than 1). The correlation suggests that containing at least 25 ruler words might be a bottom line that can be used to preclude some of the reference texts from the automated analysis.

Chart 2: amount of ruler words in each text (frequency distribution)



Source: author

Chart 3: correlation between number of measurable words and the difference between two methods



Source: author

In order to make the texts easier to reach the '25 words' bottom line, more ruler words should be included. Some researches that apply the 'training + auto evaluation' automated content analysis approach (Hopkins & King 2010, Laver 2003) include very large set of ruler-words (more than thousands normally). As mentioned above, Chinese language adds additional difficulty to the automated content analysis; however, successful attempts are also seen in recent years (for example, King 2013) that the data is crawled and analysed through R, 'readme' and other types of software.

Another concern is that some of the ruler words' linkage to the function of text is not strong enough for an accurate automated analysis to individual cases. The nature of Chinese language is that the words' meaning might 'drift'. In this study, it results in the ruler-words' high standard deviation, which points out that each word's scores in training texts are highly diversified. A salient example is the word 'industry' that has the highest standard deviation among all 15 words. Although this word is normally found in the proposals focusing on professional issues that are manually given a score below 0, exceptions are by no means real. Deputy proposal No. 18 from the 1st session of the 14th PC (being used as training

text) reflects constituents' antipathy towards the soaring pork price. The word 'industry' appears in the final section of the text that advises the policy adjustments in the 'pig-breeding industry'. Proposal No. 114 introduced in the same session illustrates local people's strong dissatisfaction with the air and sound pollution resulted in the catering 'industry'. In these two cases, the word 'industry' appears in the texts that are given positive scores.

In fact, the lack of accuracy in evaluating individual case, the principle methodological challenge confronting this paper, is not uncommon for the researches taking ACA method (Grimmer & Stewart 2013). The above discussion suggests two orientations of refining the research design that could enhance the accuracy of the evaluation of individual text.

One orientation is to make ruler words more indicative. Borrowing the philosophy of stratified sampling, the expected new method in selecting ruler words starts by stratifying training texts into several groups, say, economy, legal, and societal. 10-20 words with the highest frequency is subsequently collected from each group. Another way to make the words more indicative is to include more bigram words, instead of unigram words. To use the word 'environment' as an example. Using the word 'environment protection', instead of 'environment', avoids the peril that the word 'environment' may refer to 'industrial environment' or 'environment of investment' that strongly link to deputies' professional focus.

The second orientation is to make the scoring of the ruler words more accurate. Specific tactics may be (1) to include more training texts in the manual coding stage, which is an practical, but not principle, way to 'dilute' the peril of meaning drifting mentioned above; (2) to apply K-fold cross validation method to validate the manual scoring of the texts (Stewart & Zhukov 2009).

In addition, as too little number of ruler words affect the accuracy of the automated scoring, this paper recommends '25 words bottom line' strategy, by which only those texts with more than 25 ruler words are automatically scored. While it precludes some short texts from the automated analysis, the accuracy is guaranteed.

5. Empirical findings

Based on the result from the manually scoring process that covers all texts (both the training texts and evaluating texts), some notable empirical findings are generated, which add new knowledge to pioneering quantitative research of China's deputy-voter contact and deputy responsiveness.

First, the empirical data suggests that about a half of the city level PC deputies used their proposals as a way to express public demand, which is higher than the

result of the study focusing on provincial level (31%, see Xu 2013). It also suggests an inverse correlation between the level of the PC that a deputy serves and his or her responsiveness to the voters. It is also suggested that the inverse correlation is highly likely to be the result of the deputies' social status. One of these researches focuses on the national PC (Sang & Qiu 2010), illustrating that from the NPC deputy proposals aiming to express public voice, 75% of which are introduced by the deputies without executive or party leadership position. As for the proposals introduced by the deputies with executive or party leadership position, approximately 70% of them merely aim to report their work and progress to the central decision maker, rather than report any local claim. Li's research of a provincial level case reaches at the similar conclusion (Li 2015).

In addition, although more than a half of the proposals express public voice, the fact that the highest frequency of the cases appears in the '0-0.25' interval suggests that the most commonly seen proposals still focus on the public voice of the deputies' profession and work unit, rather than their constituency. This suggests a lack of deputy-voter linkage, as the deputies' 'vision' is limited to their profession and their own experience.

'Ridiculous proposal' is another academic focus in recent years. It is featured by its 'media hype' style or empty theme. Some of them are even clashed with moral disciplines or law. Ridiculous proposals are normally produced for three reasons that (1) bureaucracy distances the deputies holding official post from the public; (2) some deputies are too busy with their own work or profession; (3) some deputies still bear the leftist doctrine in mind (e.g. a proposal to introduce compulsory registration rule to every internet-users). At the national level, each annual NPC plenary session witnessed such proposals, which is a considerable damage to the public reputation of the PCs. The manual coding process, however, does not see such proposal. All of the proposals are scored by -2 because the theme exceeds the jurisdiction of the city level PC or cannot be handled by a specific government department, the contents are not 'ridiculous'.

6. Deputy responsiveness database

As the texts used in this paper come from a single city and a short period of time, the primary contribution of the empirical analysis is to refine and discuss the method, rather than reveal any solid empirical findings. In fact, any conclusion is substantial only if it is based on a large quantity of empirical data; once the amount of data is 'big' enough, it produces a deputy responsiveness database for empirical research.

The automated analysis enables the creation of such a database within a reasonable time and at an affordable cost; however, the principle obstacle is the official's lack of wish to disclose the texts of deputy proposals. A possible reason

for the enclosure might be the fear that 'ridiculous' and 'self-interest based' proposals may further damage the PC's public reputation.

The proposed database might have huge benefits. Vertically, the panel data will help researchers depict the history and the changing patterns of deputy-voter contact. It also contributes to the attempts to observe whether deputy responsiveness is strengthened in recent decades (in fact, text data is the only access to this, as interview and questionnaire are strong in depict present, but weak in depict past). Horizontally, a provincial or even national scale full sample database will be a solid foundation for inter-regional comparative studies, from which a series of studies can be carried out. These include the select and the case study to the 'outliner' cases (very active or passive deputy-voter connection), or a quantitative study illustrating the impact of a series of factors, say, economic development, voter's professional and education background, on the deputy-voter connections.

Reference

- Cai D. *The People's Congress of China (Zhongguo renmin daibiao dahui zhidu)*. Beijing: Law Press, 2003: 171-8.
- Cho Y. Public Supervisors and Relectors: Role Fulfillment of the Chinese People's Congress in the Market Socialist Era. *Development and Society*, 2003 (2): 197-227.
- Grimmer J, Stewart B. Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts. *Political Analysis*, 2013 (21): 1-31.
- He J. and Wang W. Deputies' Social Background and the Performance: a fsQCA Study of 13 County PCs in Beijing City (Daibiao jiegou yu lvzhi jixiao: dui beijingshi shisange quxian renda de mohuji fenxi). *Nanjing Social Science (Nanjing shehui kexue)*, 2012, (1): 78-83.
- Hopkins D, King G. A Method of Automated Nonparametric Content Analysis for Social Science. *American Journal of Political Science*, 2010 (1): 229-247.
- Huang D. and Chen C. The Performance of County People's Congress Deputies: Who is the activist? (Xianji renda daibiao lvzhi: shui geng jijì). *Sociological Studies (shehuixue yanjiu)*, 2015 (4): 169-193.
- King G *et al.* How Censorship in China Allows Government Criticism but Silences Collective Expression [J]. *American Political Science Review*, 2013 (107): 1-18.
- Laver M *et al.* "Extracting Policy Positions from Political Texts Using Words as Data [J]", *American Political Science Review*, 2003, 97.
- Li X. The Politics of 'Deistribution' in the Behaviour of People's Congress Deputies (Zhongguo rendadaibiao xingdong zhong de fenpei zhengzhi). *Open Times (Kaifang shidai)*: 2015 (4): 147.
- Manion M. Authoritarian Parochialism: Local Congressional Representation in China, *The China Quarterly*, 2014 (2): 311-338.
- Manion M. 'Good Types' in Authoritarian Elections: The Selectorial Connection in Chinese Local Congresses, *Comparative Political Studies*, 2014 (10): 1-33.
- O'Brien K. Agents and Remonstrators: Role Accumulation by Chinese People's Congress Deputies. *The China Quarterly*, 1994 (138): 359-380.
- Peng Z. *Media Content Analysis (Meijie neirong fenxifa)*. Beijing: Renmin University Press (Renmin daxue chubanshe), 2012: 13.

- Pitkin H. *The Concept of Representation*. Berkeley: University of California Press, 1967:168-9.
- Quinn K *et al.* How to Analyze Political Attention with Minimal Assumptions and Costs. *American Journal of Political Science*, 2010 (4): 209-228.
- Sang Y. and Qiu J. Evaluating PC Deputies' Performance and Social background by Formal Bills and Proposals(Cong daibiao yan he jianyi kan daibiao shuxing jiqi lvzhi zhi xiaolv). *Journal of Jiangsu Administration College (Jiangsu xingzhengxueyuan xuebao)*, 2010 (1): 76-83.
- Sieberer U *et al.* Mapping and Explaining Parliamentary Rule Changes in Europe: A Research Program. *Legislative Studies Quarterly*, 2016 (1): 61-88.
- Simon A, Xenos M. Dimensional Reduction of Word-Frequency Data as a Substitute for Intersubjective Content Analysis, *Political Analysis*, 2004 (12): 63-75.
- Stewart B, Zhukov Y, Use of force and Civil-Military Relations in Russia: An Automated Content Analysis [J]. *Small Wars & Insurgencies*, 2009 (2): 319-343.
- Truex R. *Representation within Bounds: Politics and Policy in China's National People's Congress*. Ph.D. Thesis: Yale University, 2014.
- Xu L. Understanding China's People's Congress Electoral Reform from Deputies' Proposals (Cong daibiao jianyi kan woguo rendadaibiao xuanjuzhidu de gaige). *Journal of Jinan University (Social Science Edition) (Jinan daxue xuebao, shehui kexue ban)*, 2013 (4): 36-41.
- Zou P. Study on the Full Time Representation of the People's Congress Deputies (renda daibiao zhuanzhihua wenti yanjiu). *Fudan Political Science Review (fudan zhengzhixue pinglun)*, 2008 (1): 52-79.

Biographical note:

Xingjian Yan, PhD in Politics, research fellow at Political Science Institute, East China University of Political Science and Law. The author works on comparative legislatures, especially the People's Congress institutions in China and the UK Parliament. The author currently leads a national social science fund project on promoting the public connection to the People's Congress.

Contact:

Room 226, Shixun Building
East China University of Political Science and Law
555 Longyuan Rd. Songjiang District,
Shanghai, China, 201620
yanxingj@hotmail.com