THEORIES OF DEBATES IN PARLIAMENT

Multivariate Analysis Final Data Essay

HWS 2017

Presented by:

Bengi Koseoglu

submitted to the

Political Science Group in University of Mannheim

Prof. Dr. Thomas Gschwend

Marcel Neunhoeffer

Sebastian Sternberg

December 2017

1. Purpose

Debating and voting are the most important activities of members serving in democratically elected parties. The purpose of this data essay to examine the effect of the participation of Members of Parliamentary debates in different electoral regimes, by testing two hypothesis related with ideological distances.

Hypothesis 1: In majoritarian systems, the larger the distance between MP's ideological logical position to the party, the more speeches the MP delivers in the paliament
Hypothesis 2: In proportional systems, the larger the distance between MP's ideological logical position to the party, the fewer speeches the MP delivers in the paliament

In order to test these two hypothesis, two different datasets are used, that represents Germany and England. For simplicity, U.K. assumed to have proportional regime whereas Germany assumed to have majoritarian regime, even though in Germany different electoral vote systems are present for list and district MP's.

2. Description of Datasets

2.1. United Kingdom

UK dataset consists 2 interval and dummy variables. These two interval variables include; number of speeches delivered by MP's in UK parliament between 2005 and 2009 and ideological distance which represents the MP's ideological position compare to party leader. Two dummy variables include; party variable, which indicates whether the MP is a member of conservative party and party leader variable, which shows whether the MP is the party leader. At below you may see the summary statistics of these variables and their distribution in the UK dataset.

	number_speeches	ideological_distance	conservative_MP	party_leader
min	0	0.3333	0	0
1st Quartile	67	0.6667	0	0
Median	120	1.6667	0	0
Mean	155.7	1.9489	0.3314	0.04142
3rd Quartile	204	2.4554	1	0
Max	711	5.4554	1	1
NA	0	16	0	0

Table 1: Summary Statistics for UK

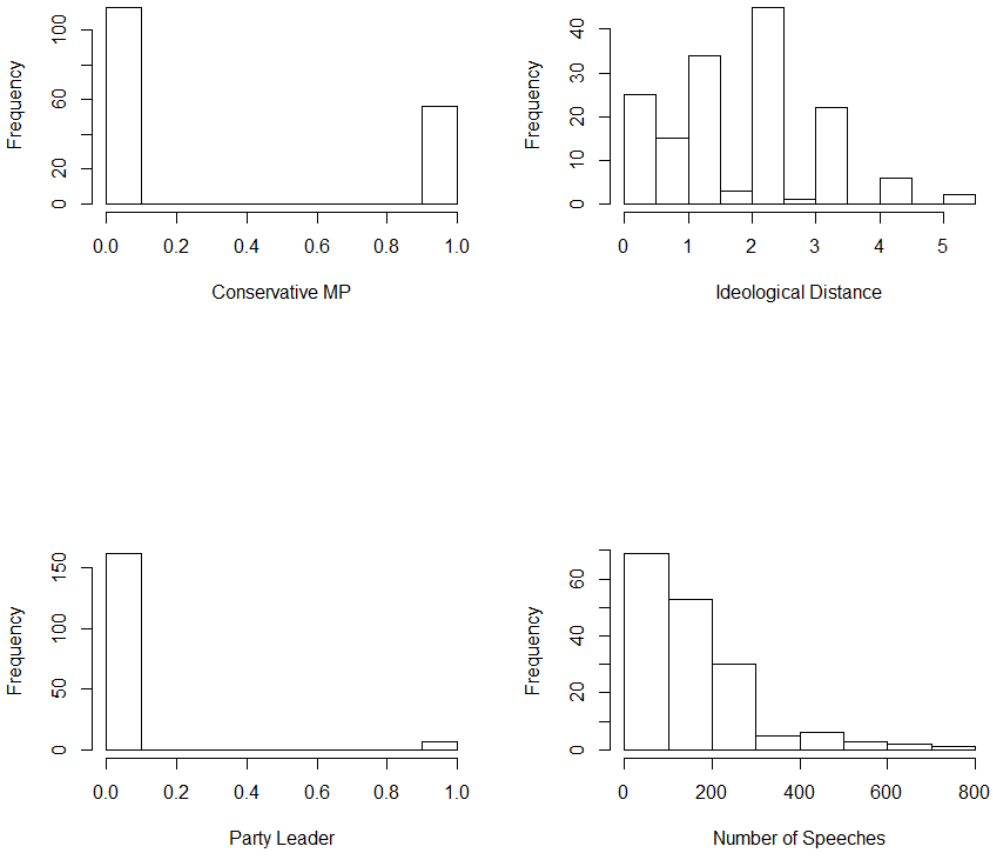


Figure 2: Histograms of Variables in UK

By looking at summary statistics and histograms, the following conclusions can be reached:

- Idealogical distance have 16 missing values
- There are more labour party members than conservatives
- There are less party leaders

- On average, in UK 155.7 speeches are given in the parliament

Let's examine, how these variables interact with our independent variable, number of speeches with aggregate functions and scatter plots.

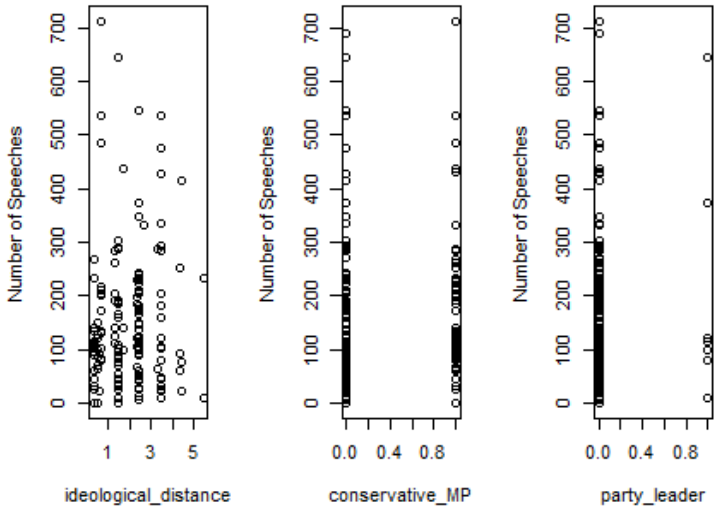


Figure 3: Scatter plot for variables in UK

By using, aggregate functions and scatter plots, following results are reached:

- Conservative party MP's make 182 speeches on average, whereas labour party MP's make 142
- Party leaders make 205 speeches on average, whereas backbenchers make 153
- Number of speeches increases as the ideological distance increases

2.2. Germany

Germany dataset consists of 3 interval, 2 dummy and 2 categorical variables that needs to be dummy encoded. These interval variables include; number of speeches delivered by MP's in Bundestag between 2005 and 2009, MP's committee assignments, and ideological distance. Dummy variables include; party leader and coalition variable which indicates whether MP is located outside of the governing coalition of SPD and CDU. Categorical variables include; MP's party and whether they are list or district. At below the summary statistics of these variables and their distribution can be found.

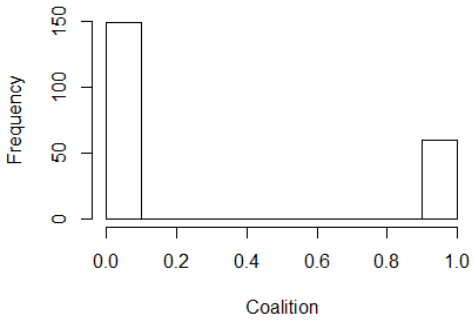
	number_speeches	ideological_distance	committee	caolMPoutside	party_leader
min	0	0	0	0	0
1st Quartile	7	0.0475	2	0	0
Median	15	0.1253	3	0	0
Mean	19.18	0.1753	0.3314	0.2871	0.09091
3rd Quartile	28	0.2603	3.196	1	0
Max	68	0.8326	4	1	1
NA	0	12	7	0	0

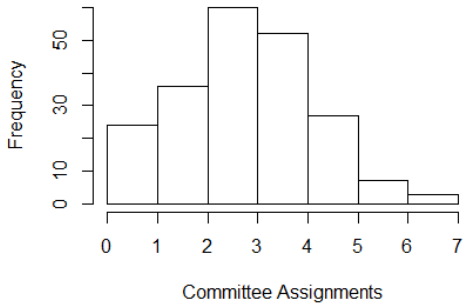
Table 2: Summary Statistics of Dummy Coded and Interval Variables for Germany

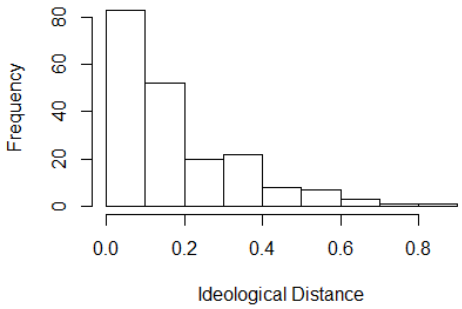
	party affiliation
FDP	28
Alliance 90/The Greens	23
CDU/CSU	60
Die Linke	21
SPD	77

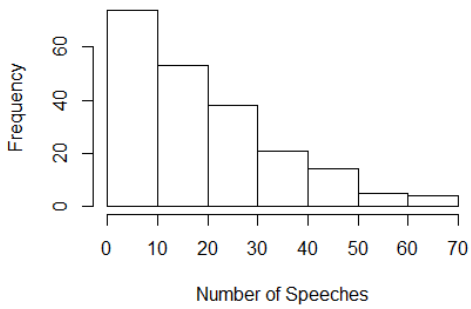
	list_candidate
District	85
List	124

Table 3: Categorical Variables in Germany


Frequency
0
50
100
150
0.0
0.2
0.4
0.6
0.8
1.0
Coalition


Frequency
0
10
30
50
0
1
2
3
4
5
6
7
Committee Assignments


Frequency
0
20
40
60
80
0.0
0.2
0.4
0.6
0.8
Ideological Distance


Frequency
0
20
40
60
0
10
20
30
40
50
60
70
Number of Speeches

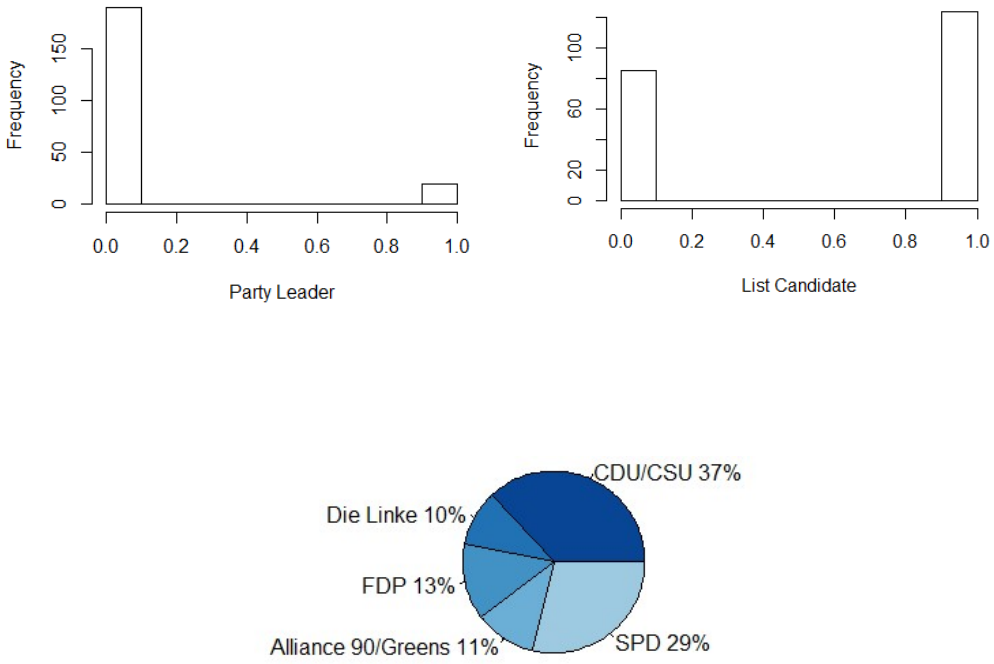


Figure 7: Histogram Plots of Variables in Germany

By looking at summary statics and histogram for Germany, the following things could be said:

- Idealogical distance variable have 12 missing values
- More MP's are outside of the coalition interval than inside
- Majority of the MP's are members of SPD and CSU/CDU
- On average, in Germany 19.18 speeches are given in the parliament, this is lower compare to UK
- There are more list candidates

Examination of these variables with number of speeches are follows.

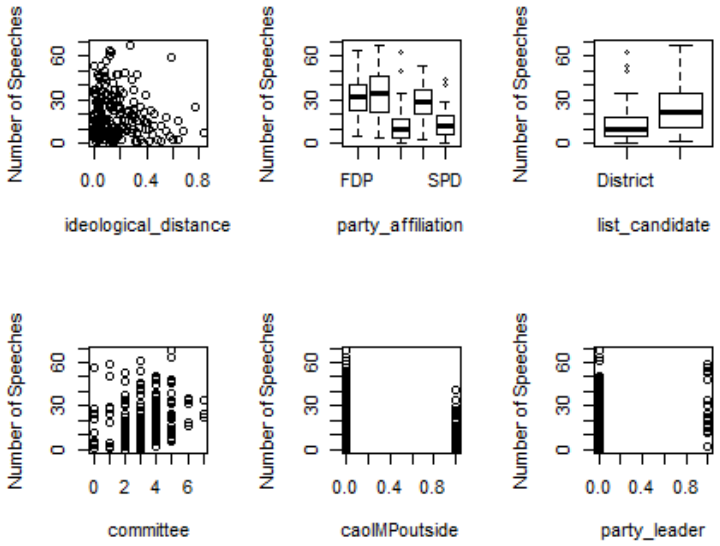


Figure 8: Scatter plot for variables in Germany

By using, aggregate functions and scatter plots, the following results are reached:

- Party leaders on average give 27 speeches, whereas backbenchers give 18
- List candidates tend to give more speeches with 23 speeches
- Higher committee assignment leads to higher speeches
- MP's inside the coalition give more speeches compared to MP's outside the coalition
- MP's in Green party gave the highest speeches followed by FDP, Die Linke, SPD, CDU/CSU

3. Modeling

3.1. Majoritian Systems: U.K.

Missing Value Handling: For handling missing values in ideological distance variable, first these missing values are examined closer for patterns but no pattern has been detected. Therefore, it was decided to fill missing values with their median value due to the fact that median is more robust to outliers than mean statistics.

Outliers Handling: For outliers handling, box plots of number of speeches and ideological distance variables are examined.

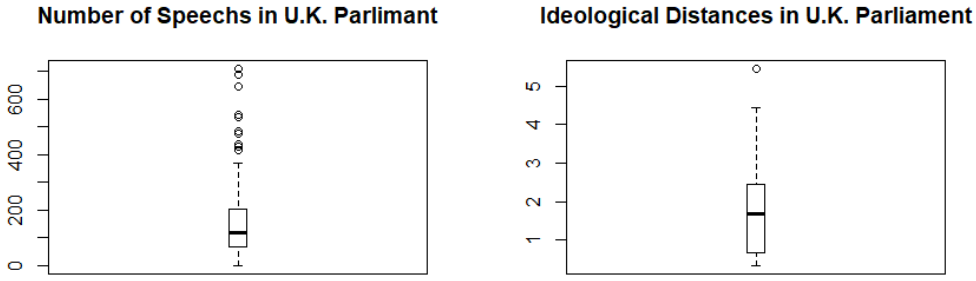


Figure 9: Box Plot of Interval Variables in UK

As understood from the graph, both of these variables have observations that are different from the rest of the population. Having outliers in the model can be problematic since they can effect the accuracy of models. On

the other hand, these observations may contain important information regarding number of speeches in UK parliament. Therefore, a new dataset is created by removing observations that lie outside 1.5 Quartile Range and models are built on both datasets. After modeling, the best dataset is determined by Akaike Information Criterion(ACI).

Modeling: For modeling, negative binomial is used, since there are overdispersion in the dataset due to having larger variance in response variable than its mean. The following models are built on two different datasets using the same settings.

	Dependent variable:			
	number_speeches			
	(1)	(2)	(3)	(4)
party_leader	0.360		0.296	0.427
	(0.339)		(0.342)	(0.335)
conservative_MP	0.280*	0.262*		0.464***
	(0.147)	(0.148)		(0.163)
ideological_distance			0.062	0.180**
			(0.071)	(0.078)
Constant	4.932***	4.957***	4.924***	4.537***
	(0.086)	(0.085)	(0.147)	(0.185)
Observations	161	161	161	161
Log Likelihood	−970.960	−971.577	−972.419	−968.610
θ	1.307*** (0.137)	1.298*** (0.136)	1.286*** (0.135)	1.341*** (0.142)
Akaike Inf. Crit.	1,947.919	1,947.154	1,950.837	1,945.219

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

Table 4: Negative Binomial Models on Outliers Handled Dataset in UK

	Dependent variable:			
	number_speeches			
	(1)	(2)	(3)	(4)
party_leader	0.357		0.301	0.413
	(0.341)		(0.345)	(0.340)
conservative_MP	0.262*	0.244*		0.379**
	(0.145)	(0.145)		(0.154)
ideological_distance			0.038	0.114*
			(0.060)	(0.064)
Constant	4.937***	4.960***	4.960***	4.670***
	(0.085)	(0.084)	(0.136)	(0.166)
Observations	169	169	169	169
Log Likelihood	−1,019.123	−1,019.725	−1,020.585	−1,017.800
θ	1.287*** (0.132)	1.279*** (0.131)	1.268*** (0.129)	1.305*** (0.134)
Akaike Inf. Crit.	2,044.245	2,043.450	2,047.170	2,043.600
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 5: Negative Binomial Models on Outliers Present Dataset in UK

As can be seen from the built models, in terms of variable significance, the models yields similar results both on outlier handled and outlier present

datasets. However, from AIC perspective, outlier handled dataset have lower values, which indicates a better fit of model. Therefore, model interpretations are done on outlier handled dataset. First, a base model is created while excluding ideological distance in order to see the combination of effect of independent variables on number of speeches. In this model, party variable is significant and have a positive effect on the dependent variable. When the effect of party leader is removed , party variable remained significant and positive(model2). Then, the effect of ideological distance controlling for party leader is examined, no significance effect has seen (model3). When party variable is included to the model3, ideological distance became significant with a positive effect. Among the built models, model4 that was built on the outlier handled dataset while including all the variables, chosen as the best model due to having lowest AIC.

Simulations and First Differences: In order to understand the magnitude of the effects and associated uncertainty from model4, several simulations are done. First simulation is done, to see the impact of different ideological distances of backbenchers on conservative and labour parties.

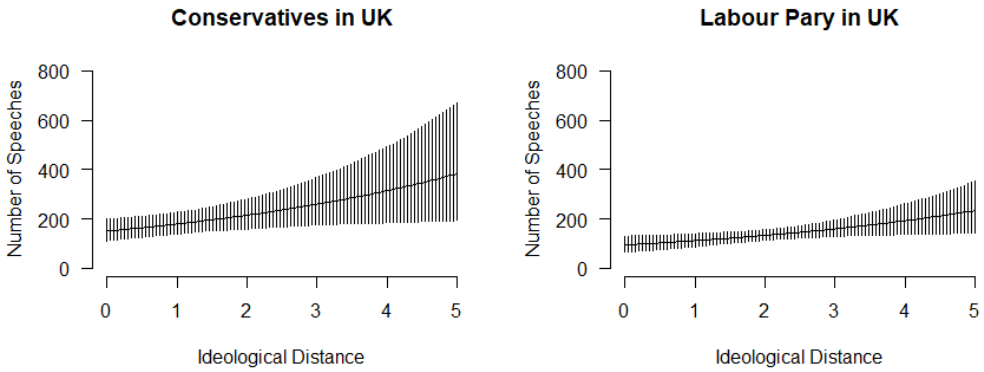


Figure 10: Simulation of Ideological distance in Conservative and Labour Parties in U.K.

By looking at the graphs it can be concluded that as the ideological distance increases, MP's speak more frequently in UK, regardless of the party. Therefore, we can conclude that in majoritarian systems, the larger the distance between an MP's, ideological position to the party, the more speeches the MP delivers in parliament and accept our hypotheses for majoritarian systems. Also, in Conservative and Labour parties, the uncertainty of our estimation increases as ideological distance increases, this uncertainty is higher for conservative party. Furthermore, in order to compare these two parties in terms of the number of speeches given by their backbenchers, first differences between Labour and Conservative parties are examined while holding ideological distance at median.

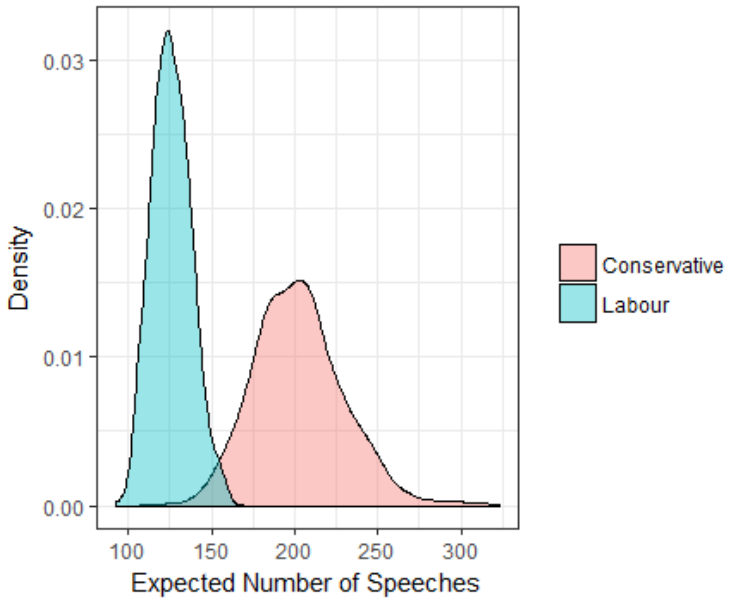


Figure 11: First Difference of Expected Number of Speeches in Conservative and Labour Parties U.K.

From the first difference analysis, it was observed that conservative MP's on average give more speeches compare to their peers in Labour party. Lastly, to understand the effect of party leaders on party level in majoritarian systems, first differences of party leaders are examined on party level.

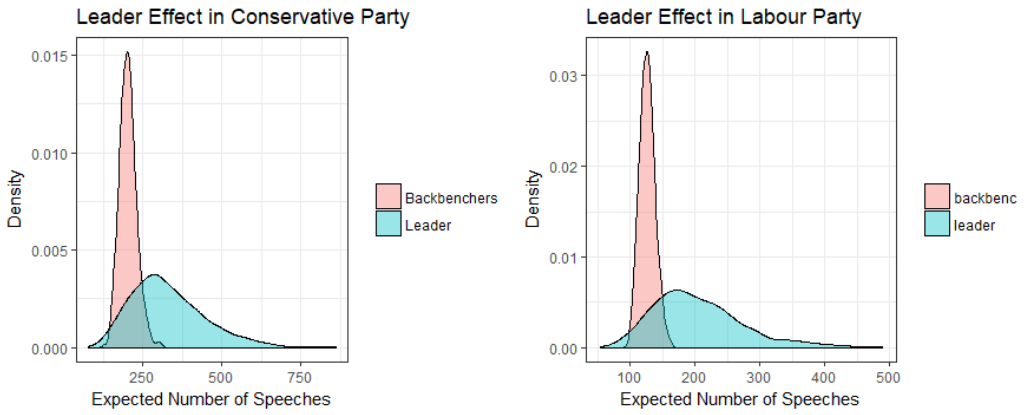


Figure 12: First Difference of Expected Number of Speeches in Conservative and Labour Parties U.K. in terms of Party Leaders

For both of the parties, party leaders expected to give more speeches. Therefore, the effect of the party leader is robust to different parties.

3.2. Proportional Systems: Germany

Missing Value Handling: The missing values in Germany dataset are examined for patterns and no pattern regarding their missingness have been found. Therefore missing values are filled with their median value.

Outlier Handling: After handling missing values, interval variables in the UK dataset have been examined for outliers using box plots.

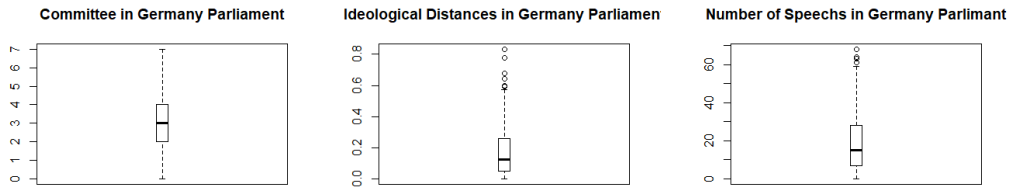


Figure 13: Box Plot of Interval Variables in Germany

As can be seen from the box plots, committee variable doesn't show any potential outliers, whereas number of speeches and ideological distance variables show some potential outliers. However, no outlier handling is done, since the outliers in number of speeches are not deviating from the population as much as UK and scale of ideological distance is between 0 and 1.

Data Preparation: Party affiliation and list candidate variables are dummy coded. Furthermore, new variable is added to the dataset that represents the number of seats that the parties received in 2005 elections. Because, in proportional systems, the higher the seatshare of a party, the more the party wants to hold their unity and therefore as the seatshare of a party and MP's distance increases, MP's are expected to take less floor time.

Modeling: In modeling, negative binomial is used due to overdispersion

in the data. First a base model is built, while excluding the ideological distance and party dummy variables, in order to see the committee assignment, party list and party leader's combination effect. From the base model, it was observed that party leaders, list candidates and MP's with higher committee assignments speak more often. When ideological distance is introduced to base model, the effect was negative, which supported our hypothesis regarding proportional systems. The negative effect of ideological distance continued to hold when it was controlled for the different parties and governing coalition (model 4).To understand the effect of governing coalition, a new model is built by keeping CDU/CSU and SPD dummy party variables (model5). In the model, the effect of coalition parties and ideological distance remained statistical significant and negative. To understand the impact of each party on number of speeches, four more models are built by including different party dummies(model 6-9). Through these models, the significance of ideological distance disappeared but remained negative.

	Dependent variable:								
	number_speeches								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
ideological_distance		−0.589*	−0.461	−0.564*	−0.524*	−0.410	−0.477	−0.390	−0.417
		(0.325)	(0.301)	(0.315)	(0.315)	(0.336)	(0.336)	(0.336)	(0.341)
party_leader	0.502***	0.502***	0.278*	0.289*	0.270	0.425**	0.417**	0.488***	0.430**
	(0.178)	(0.177)	(0.167)	(0.166)	(0.165)	(0.175)	(0.175)	(0.175)	(0.181)
list_candidate	0.621***	0.611***	0.098	0.099	0.107	0.512***	0.415***	0.476***	0.544***
	(0.107)	(0.107)	(0.121)	(0.120)	(0.121)	(0.109)	(0.113)	(0.112)	(0.111)
committee	0.145***	0.146***	0.124***	0.127***	0.127***	0.141***	0.137***	0.137***	0.141***
	(0.035)	(0.035)	(0.032)	(0.032)	(0.032)	(0.034)	(0.034)	(0.035)	(0.035)
FDP			−0.085	−0.091				0.366**	
			(0.185)	(0.185)				(0.157)	
Die.Linke			−0.254	−0.261					0.135
			(0.202)	(0.201)					(0.179)
CDU.CSU			−0.943***	−0.984***	−0.873***		−0.324***		
			(0.183)	(0.187)	(0.149)		(0.120)		
SPD			−0.869***	−0.944***	−0.834***	−0.315***			
			(0.175)	(0.187)	(0.149)	(0.120)			
caolMPoutside				0.137	0.131	−0.040	−0.215*	−0.142	−0.179
				(0.130)	(0.130)	(0.135)	(0.125)	(0.127)	(0.128)
Constant	2.003***	2.103***	3.057***	3.066***	2.947***	2.270***	2.380***	2.158***	2.168***
	(0.140)	(0.154)	(0.229)	(0.229)	(0.195)	(0.164)	(0.171)	(0.158)	(0.160)
Observations	209	209	209	209	209	209	209	209	209
Log Likelihood	−799.550	−798.049	−779.188	−778.646	−779.496	−793.968	−794.028	−794.181	−796.705
θ	1.979*** (0.215)	2.008*** (0.219)	2.489*** (0.288)	2.502*** (0.290)	2.477*** (0.286)	2.100*** (0.232)	2.104*** (0.233)	2.097*** (0.231)	2.039*** (0.223)
Akaike Inf. Crit.	1,607.100	1,606.099	1,576.376	1,577.291	1,574.991	1,601.935	1,602.057	1,602.363	1,607.410
Note:									*p<0.1; **p<0.05; ***p<0.01

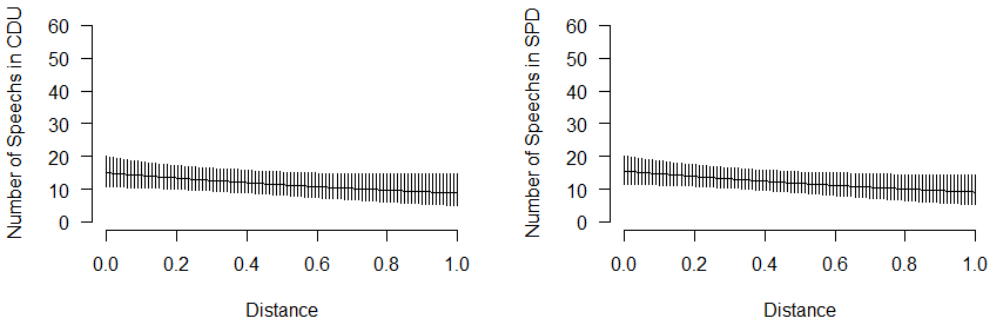
Table 6: Negative Binomial Models on Germany with Existed Variables

Furthermore, a new base model is created by adding seatshares, while controlling for party leader, committee and coalition (model10). In the base model, the variable is found highly significant with negative but small impact. This held when ideological distance was also added (model11). However, when party dummies were added, the variable lost significance and resulted with missing values,due to multicollinearity with party dummies. Among built 12 models, model 4 is chosen as the best model due to having more explanatory power.

	Dependent variable:		
	number_speeches		
	(10)	(11)	(12)
ideological_distance		−0.532*	−0.564*
		(0.315)	(0.315)
party_leader	0.236	0.249	0.289*
	(0.166)	(0.165)	(0.166)
list_candidate			0.099
			(0.120)
committee	0.124***	0.126***	0.127***
	(0.033)	(0.032)	(0.032)
FDP			−0.091
			(0.185)
Die.Linke			−0.261
			(0.201)
CDU.CSU			−0.984***
			(0.187)
SPD			−0.944***
			(0.187)
numberofseats	−0.005***	−0.005***	NA
	(0.001)	(0.001)	
caolMPoutside	0.072	0.141	0.137
	(0.121)	(0.127)	(0.130)
Constant	3.290***	3.367***	3.066***
	(0.163)	(0.170)	(0.229)
Observations	209	209	209
Log Likelihood	−781.293	−779.918	−778.646
θ	2.434*** (0.281)	2.467*** (0.285)	2.502*** (0.290)
Akaike Inf. Crit.	1,572.586	1,571.836	1,577.291
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 7: Negative Binomial Models on Germany with New Added Variables

Simulations and First Differences: In order to understand the magnitude of the effects and associated uncertainty found from the model4, simulations are made. Firstly, to understand the effect of ideological distance on number of speeches given by backbenchers outside coalition on party level, simulations are conducted for each party, by simulating ideological difference and holding list candidate and committee on median. The simulations indicates that, the more the MP's ideological distance deviates from party, the less the MP's take floor regardless of the parties. Therefore, we can conclude that in proportional systems, the larger the distance between an MP's, ideological position to the party, the fewer speeches the MP delivers in parliament and this is opposite of what have been observed in UK. Also, we are more certain about number of speech estimates in CDU and SPD due to narrower confidence intervals.



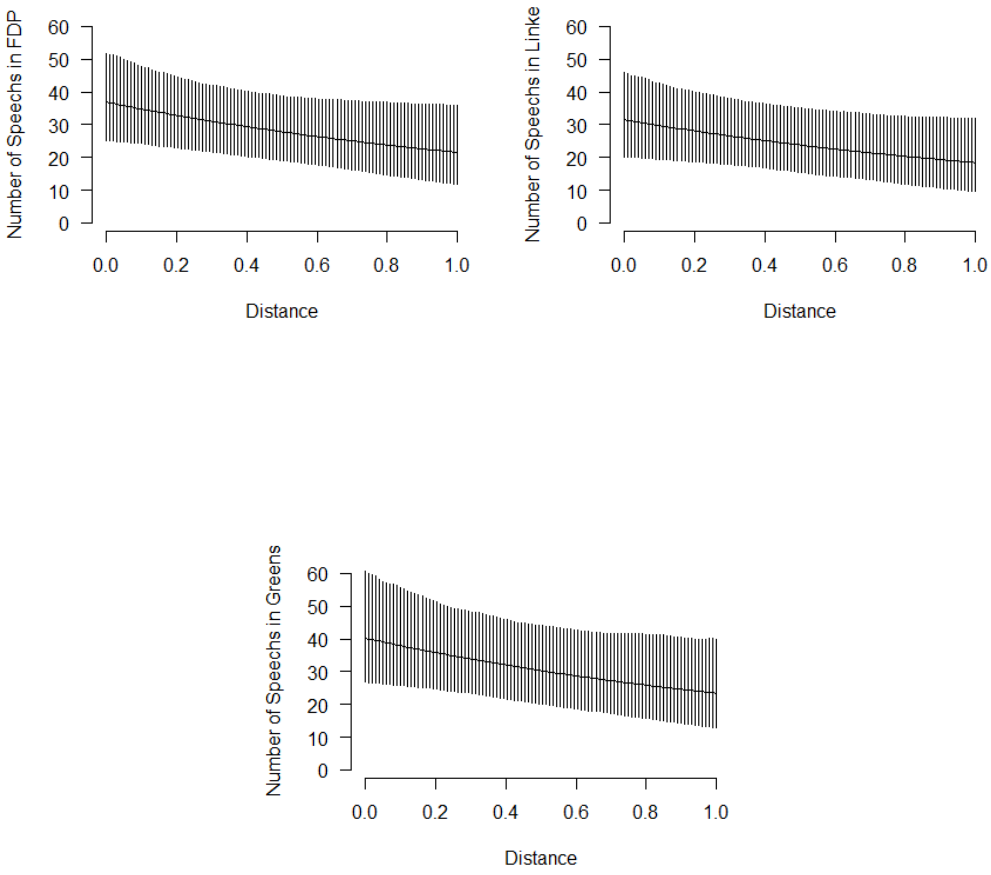


Figure 16: Simulation on Ideological Distance with Different Party Levels

Furthermore, to understand the differences in number of speeches given by backbenchers outside coalition interval in each party, first differences are examined by holding ideological distance, list and committee assignments at their median. According to the graph, CDU/CSU and SPD MP's are expected to give around 15 speeches, whereas DieLinke, FDP and Greens are expected give 25, 30 and 35.

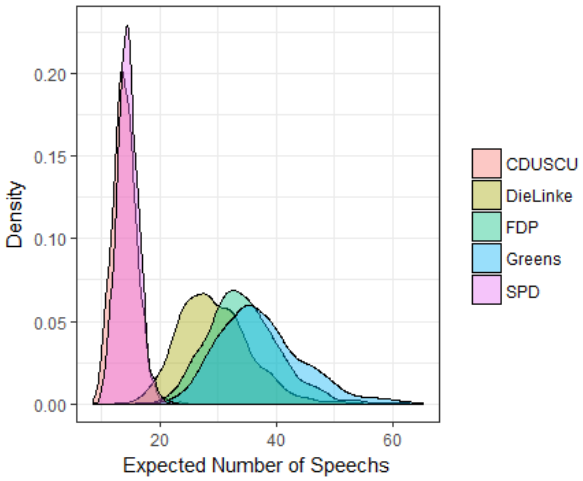
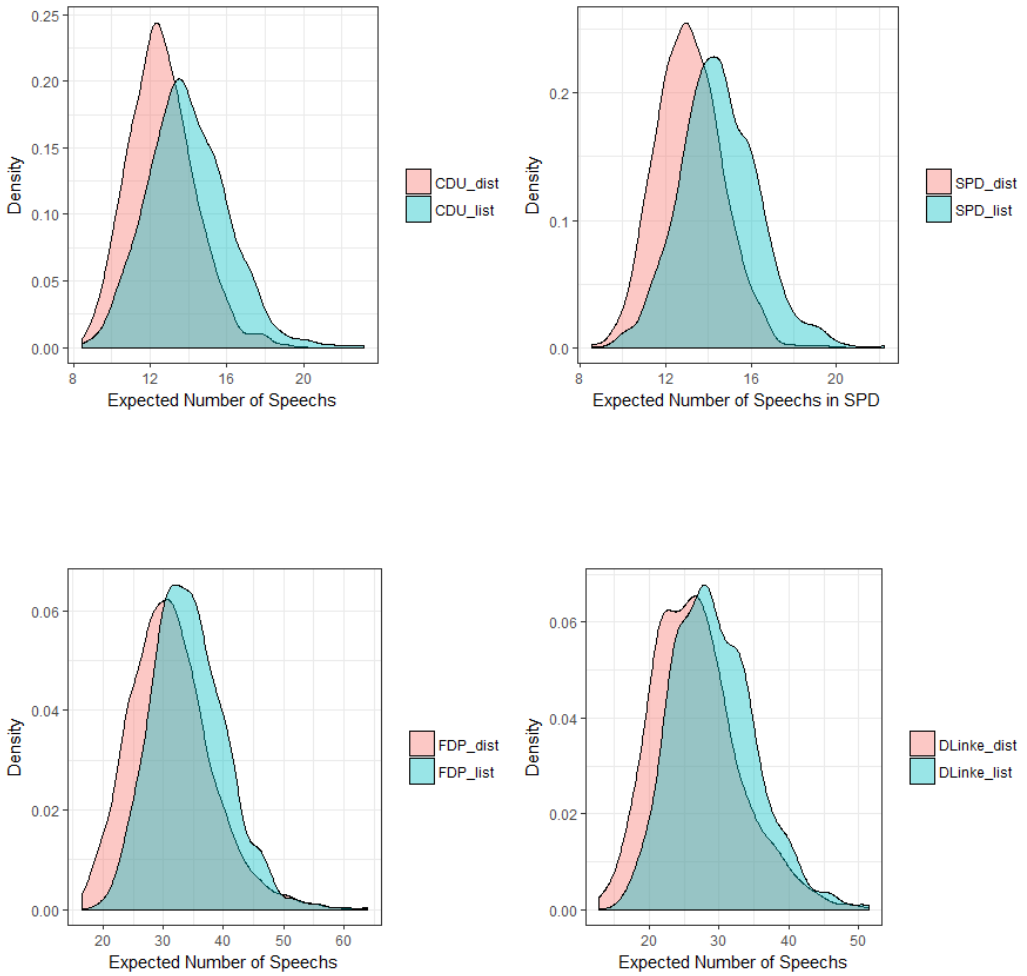


Figure 17: First Difference of Expected Number of Speeches in Parties in Germany

Morever, to understand the effect of list and district MP's on number of speeches given by backbenchers outside coalition in each party, first differences are examined by holding ideological distance and committee assignments at median. Through paper,it was assumed that Germany have proportional electoral system, even though Germany has mixed proportional system. During modeling phase, list variable is added to the model as control variable in order to take list and district differences into account. As a result of first difference analysis, no major differences of speeches regarding list and district MP's are found on party level.


Density
0.25
0.20
0.15
0.10
0.05
0.00
8
12
16
20
Expected Number of Speechs
CDU_dist
CDU_list
Density
0.2
0.1
0.0
8
12
16
20
Expected Number of Speechs in SPD
SPD_dist
SPD_list
Density
0.06
0.04
0.02
0.00
20
30
40
50
60
Expected Number of Speechs
FDP_dist
FDP_list
Density
0.06
0.04
0.02
0.00
20
30
40
50
Expected Number of Speechs
DLinke_dist
DLinke_list

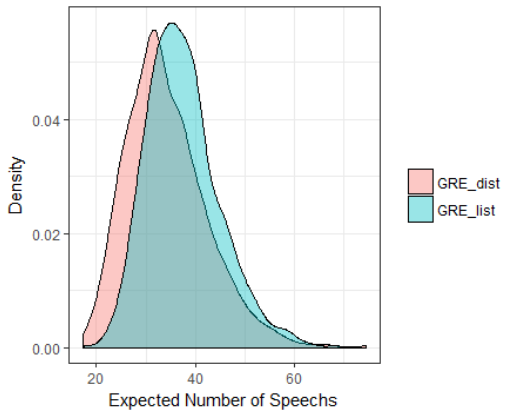


Figure 20: First Difference of List and District MP's on Number of Speeches in Parties in Germany

Lastly, another simulation is made in order to see the effect of seatshare on number of speeches made by backbenchers outside coalition while holding ideological distance, committee on median by using model11. As can be seen from the graph, our intuition was right, as MP's party's seat shares increases, the less speeches MP's made.

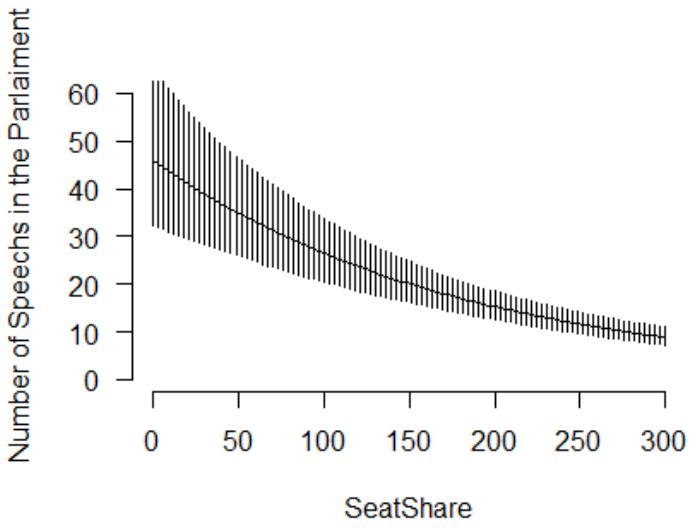


Figure 21: First Difference of Expected Number of Speeches in Conservative and Labour Parties U.K.

4. Conclusion

To conclude, in this proposal the effect of participation of MP's to Parliamentary debates in different regimes is examined by testing two hypothesis, while assuming Germany as proportional and UK as majoritarian. The fact of Germany having two different electoral vote systems in district and list level is controlled by adding list as an independent variable to models. To test the hypothesis for majoritarian systems, a simulation is made on party level on UK dataset and was found that the larger the distance between MP's ideological position to the party, the more speeches the MP delivers in the parlaiment regardless of the party, therefore hypothesis is accepted. In Ger-

many, the hypothesis about proportional systems is tested with a simulation on party level, and it was found that the larger the distance between MP's ideological position to the party, the fewer speeches MP delivers in parlaiment, this effect held on party level, therefore hypothesis 2 is accepted. In order to explain the differences more, several first difference and simulations are done, details can be found in the paper.

For further research, MP's committee's such as budget committee, can be taken into account by adding new variables on MP level. Because members of different committee's can be make different number of speeches. Moreover, by examining more countries regarding parliamentary debates, our hypotheses can be more generalizable.

(Word Count:2198)

Bibliography

[1] @Book, author = Hadley Wickham, title = ggplot2: Elegant Graphics for Data Analysis, publisher = Springer-Verlag New York, year = 2009, isbn = 978-0-387-98140-6, url = http://ggplot2.org,

[2] @Manual, title = stargazer: Well-Formatted Regression and Summary Statistics Tables, author = Marek Hlavac, year = 2015, note = R package version 5.2, organization = Harvard University, address = Cambridge, USA, url = http://CRAN.R-project.org/package=stargazer,

[3] @Book, title = Modern Applied Statistics with S, author = W. N. Venables and B. D. Ripley, publisher = Springer, edition = Fourth, address = New York, year = 2002, note = ISBN 0-387-95457-0, url = http://www.stats.ox.ac.uk/pub/MASS4,