Supporting Information (SI)

Supp Info A Pseudo-code for burstiness calculation

Let `tdm` by the term-document matrix of the speeches, such that each row is a word, and each column a speech. A given i^{th}, j^{th} cell-entry of `tdm` is a binary indicator $\{0, 1\}$ of whether or not word i appeared in speech j (multiple uses are treated similarly to single occurrences). The steps to calculate our statistics are as follows:

```
for (i in 1:number of rows in tdm){
```

1. draw the ith row of `tdm`, which is a binary vector of occurrences. Thus, supposing there were 10 speeches in the corpus, we might have {`0,0,1,1,1,1,0,0,0,0`}, with the use of some term appearing in documents 3,4,5,6 and then nowhere else.

2. calculate the burstiness of this term, as described above. That is, for each individual burst, multiply its level (via the estimated i) by its duration (literally, how long this particular value of i lasts until a new value of i arises). Then sum these terms for all bursts that occur for the term. Denote this sum as `b`.

3. allocate `b` to the appropriate location in a document vector (that is, a vector of length equal to the number of documents in the corpus). Notice that this will require simply adding it to whatever the 'running total' for that document currently is (since a given document may have multiple bursty terms).

4. record the time of the start of the maximum or 'peak' burst of the ith term, and the time of the end of that burst. Using a look-up table, record the MP making the speech that began the peak burst.

}

The result of this algorithm is (a) a table of bursty terms (i.e. all terms with non-zero burstiness), each with a starting and ending point of their peak burst, and the identity of the MP who began the peak burst; (b) a table listing every speech and the burstiness of each (which will generally be zero for at least some speeches). Finally, a look-up table is used to aggregate the results of the speech table by MP: that is, each speech is mapped to a unique MP, and his score derived by summing the total burstiness of all the speeches he gave (some of which may be zero scored).

Supp Info B Correlation between burstiness and 'speechiness' over time

Figure 9 reports the correlation between all MPs' burstiness and the number of speeches they gave, for the various sessions in the data. Notice that the mean is around 0.6, implying that though the variables are correlated, they do not apparently measure identical concepts. For completeness, we also investigated the correlation between an MP's 'speechiness' rank in a given session, and their burstiness rank during the same period with the idea that the relative share of speeches might provide a better measure of importance within the opposition than the raw number of speeches. In reality, the correlation varies between 0.66 and 0.90, with a mean of 0.83. Although this is higher, there is clearly some room left between the measures.

Supp Info C Robustness: using 'raw' burstiness in regression

Recall that, in our account, being an 'leader' (and thus in the shadow cabinet) is associated with being an *outlier* in burstiness terms. In Table 3 we consider the robustness of this

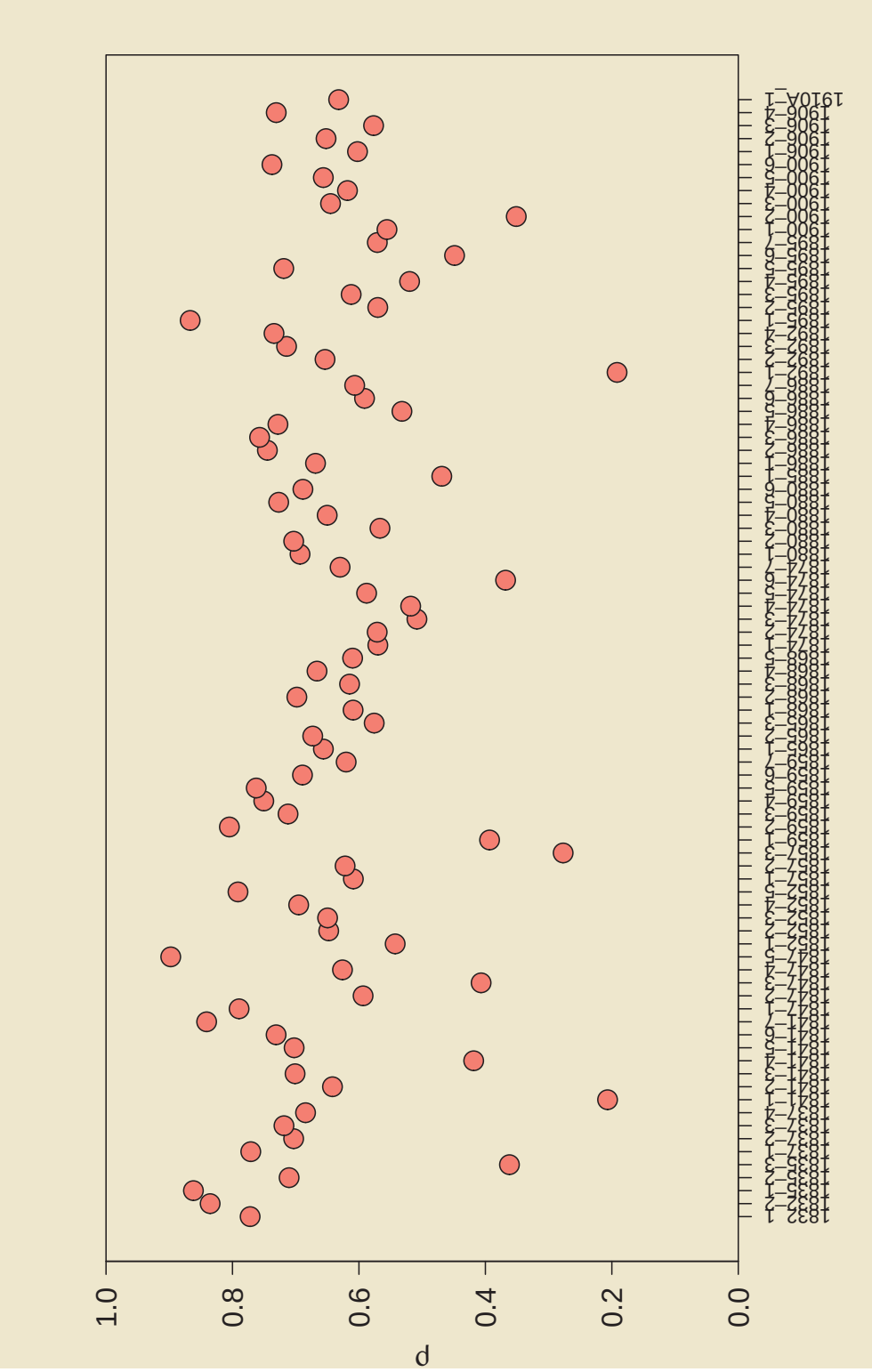

Figure 9: Speechiness is not burstiness: correlation between burstiness of opposition MPs and number of speeches they gave, over time.

	Model I	Model II
(Intercept)	-6.0205^{***}	-3.6912^{***}
	(0.5228)	(0.2739)
outlier	3.9864^{***}	
	(0.5857)	
session number	0.0364^{***}	0.0079
	(0.0090)	(0.0055)
burstiness	0.0000^{***}	$0.0000^{\dagger}$
	(0.0000)	(0.0000)
outlier×session number	-0.0358^{**}	
	(0.0117)	
burstiness×session number		-0.0000
		(0.0000)
N	3076	3076
AIC	856.9049	987.1880
BIC	977.5326	1083.6902
$\log L$	-408.4525	-477.5940

† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Table 3: Raw Burstiness as a control and alternate measure of 'leadership' on opposition benches (standard errors clustered at the MP level).

definition by fitting models using raw burstiness as an alternative measure of leadership (both additionally and as a substitute for outlier status).[27] As we see in Model I, the coefficient on outlier status is essentially unchanged, though a little extra explanatory power is added (AIC is lower). All in all, this suggests that it is being an outlier that 'matters', rather than simply being bursty from the opposition backbenches. Finally, we consider Model II that does not use our outlier status variable at all, and relies solely upon an MP's burstiness, in addition to the time variable and the interaction. Note that this model does a relatively poorer job (in terms of fit) than the variant using the outlier metric we explained above.

[27]Outlier status is obviously post-treatment with respect to raw burstiness, and so one should not be overly confident in interpreting the coefficients.