Appendix

In this Appendix, we explain how we processed and analyzed the speeches at parties' national congresses and show the results from a number of alternate model specifications for the analysis presented in "Leadership Competition and Disagreement at Party National Congresses." In the following section we describe the process for collecting and analyzing party congress speeches. In the second section, we then show the results of some secondary analyses replicating the results from Table 2 in the main text.

Text Processing and Position Analysis

We first collected the transcripts from the party congresses either through the parties' own websites or through their associated political foundations. We limited our data collection to transcripts from party congresses which contain full documentation of the speeches given to ensure that we have a full population of the scheduled speeches at these meetings. We then split the PDF documents and save the speeches of individual speakers in separate text files. We combined speeches from the same speaker at a conference in the context in which they spoke multiple times. We exclude speeches from the corpus given by honorary speakers and other non-party members. We then clean the files by filtering out all content that is not produced by the speakers themselves (introduction/announcement of the speaker, interruptions by the audience, section headings, etc.). We also transform all words to lower case and remove punctuation, numbers, additional white spaces, hyphenation, stop words such as articles or conjunctions and frequently occurring words without substantive content such as personal names or greetings. Finally, we apply the German and French Porter stemming algorithm implemented in the *R* package TM to reduce all words to their basic word stem. The processed documents are then combined into a document-term-matrix, in which rows correspond to documents and columns to word frequencies.

Before applying the WORDFISH model we have to globally identify the model. We do this by setting the mean of the positions (ω) to zero and the standard deviation to one. This strategy requires us to indicate two documents, of which the first document has a mean positions that is more negative than the second. For all party congresses, we select documents that we believe represent both ends of the ideological spectrum. We then estimate the positions of all documents using the “wordfish” function implemented in the *R* package *Austin*.

After we estimate the position of the individual speakers, we run diagnostics to ensure that the words used in the estimation contain politically relevant information and discriminate between the different ends of the ideological dimension. To illustrate this process, Figure 1 plots the word fixed effects against the word weights for the 2000 Congress of the PS in Grenoble. The word fixed effects indicate how frequently a word appears in the corpus and the weight relates to its placement on the ideological dimension. Ideally, the plot should resemble an “Eiffel Tower”, in which words with high fixed effects have a weight of zero, while words with low fixed effects have either a positive or negative weight.[1]

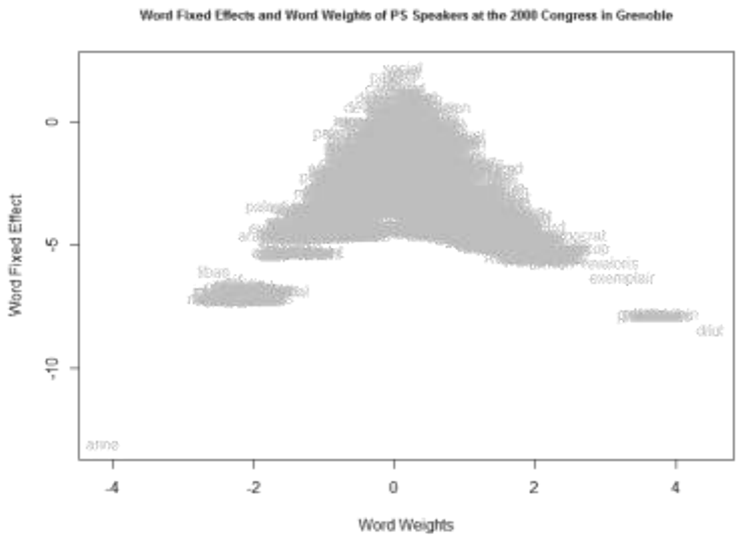

Figure 1. Word fixed effects and weights from the 2000 Grenoble Congress.

[1] see Slapin and Proksch 2008.

Figure 2 shows a more detailed diagnostics of words stems from the same party congress describing the economy. As expected, words referring to labor rights such as “solidar” (solidarity) have weights that are on the opposite side of the dimension from words that might indicate support of business interests such as “entrepris” (enterprise).

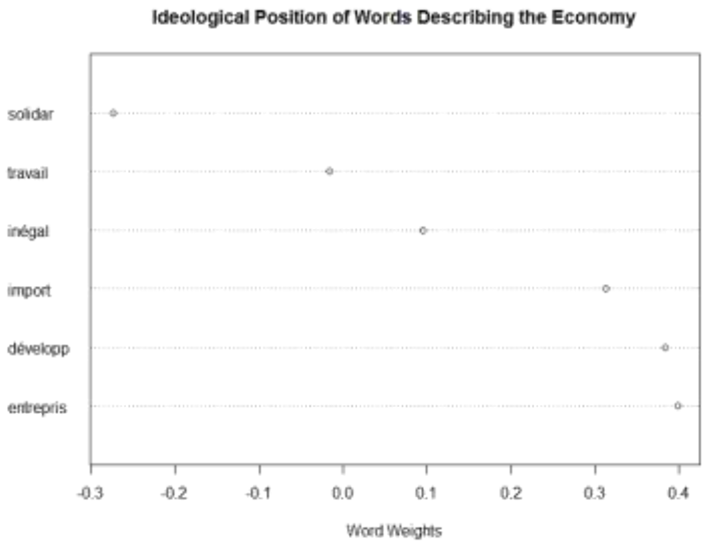

Figure 2. Selected word weights from the 2000 Grenoble Congress.

Additional Analysis

We include different modeling choices such as random effects, fixed effects or multi-level models as well as alternate specifications. While the substantive inferences drawn from the results stay largely the same, the results vary somewhat depending on the exact model specification. Full tests of these alternate explanations such as second order election effects require additional observations.

An alternate explanation that could explain changes in intra-party heterogeneity could link to simply whether or not there is an impending election. We operationalize this variable dichotomously to refer to whether an election occurred in the same year as the congress. Previous work by Ceron might lead us to expect that an election year will lead to some sort of unity effect where parties become more cohesive in that year. We also include a dummy variable to account for potential second order election effects from election held for the European Parliament. We expect that EP elections could cause parties to be less cohesive as they try out alternate strategies in those elections or more cohesive as they seek to demonstrate their unity.

The results from this analysis are presented in Table 1. Consistent with Ceron's (2012) logic, there is some evidence that there is an election year effect. The coefficient is negative and weakly significant in both Model 1 and Model 3. This indicates that parties are more cohesive in election years. However, there is less evidence of an impact from second order elections. The dummy variable for European Parliament election years is positive, but not significant in either model. Further analysis will be required to disentangle the relationship between second order elections and intra-party disagreement.

Due to the small number of observations in our analysis, our exact modeling choice could potentially influence our results. Therefore, in Table 2 we demonstrate the results of our analysis using alternate approaches to accounting for the cross-sectional time series nature of our data. In particular, we rerun the analysis using simple Huber-White robust standard errors (as in the paper), clustered standard errors (on the year), random effects for years, fixed effects for years and finally a multi-level model with random intercepts for the year and party level. While not all of these models may be fully appropriate given the small sample size, the results tell a similar story. In all except the fixed effects model, the primary independent variables are in the same direction as in the model with simple robust standard errors and are statistically significant.

Table 1. Robustness checks with EP and Parliamentary Election variables.

	Model 1	Model 2	Model 3
Government Party	0.042	0.063	0.048
	(0.123)	(0.131)	(0.126)
Δ % GDP Growth	9.162**	9.456**	9.193**
	(2.754)	(2.743)	(2.814)
Government Party X Δ % GDP Growth	-8.893*	-10.539*	-9.186*
	(3.875)	(4.397)	(4.012)
Lost Government	0.307	0.162	0.311
	(0.224)	(0.203)	(0.225)
Δ % Parliamentary Vote	-0.002	-0.010	-0.001
	(0.010)	(0.010)	(0.010)
Lost Presidency	0.009	0.008	0.009
	(0.006)	(0.006)	(0.006)
Δ % Presidential Vote	0.064	0.088	0.068
	(0.106)	(0.112)	(0.109)
PS dummy	-0.469**	-0.335*	-0.465*
	(0.172)	(0.136)	(0.175)
SPD dummy	-0.306	-0.205	-0.308
	(0.183)	(0.159)	(0.184)
UMP dummy	-0.733	-0.317	-0.754
	(0.471)	(0.396)	(0.479)
Parliamentary Election Year	**-0.344**+		**-0.350**+
	(0.172)		**(0.175)**
EP Election Year		**0.041**	**0.067**
		(0.136)	**(0.120)**
Constant	1.826***	1.704***	1.814***
	(0.141)	(0.127)	(0.149)
R^2	0.318	0.267	0.320
Root Means Squared Error	0.407	0.422	0.411
AIC	64.759	68.813	66.575
BIC	85.013	89.067	88.854
Observations	56	56	56

Table 2. Alternate Standard Errors

	Model 4 Robust SE (Model 8 in the paper)	Model 5 Clustered SE (Year)	Model 6 Random Effects (Year)	Model 7 Fixed Effects (Year)	Model 8 Multilevel Model
Government Party	0.059	0.059	0.049	0.024	0.065
	(0.128)	(0.123)	(0.119)	(0.164)	(0.127)
Δ % GDP Growth	9.434**	9.434**	9.618**	11.444	9.911***
	(2.709)	(2.735)	(3.333)	(10.018)	(2.056)
Government Party X Δ % GDP Growth	-10.346*	-10.346**	-10.565**	-7.328	-10.083***
	(4.281)	(3.697)	(3.886)	(4.791)	(2.810)
Lost Government	0.161	0.161	0.071	0.613	0.129
	(0.204)	(0.208)	(0.325)	(0.584)	(0.228)
Δ % Parliamentary Vote	-0.011	-0.011	-0.010	0.065	-0.017***
	(0.009)	(0.009)	(0.019)	(0.039)	(0.002)
Lost Presidency	0.008	0.008	0.013	0.097*	0.010***
	(0.006)	(0.007)	(0.012)	(0.042)	(0.002)
Δ % Presidential Vote	0.085	0.085	-0.514	-0.747	0.071
	(0.109)	(0.118)	(0.419)	(0.521)	(0.092)
PS dummy	-0.339*	-0.339*	-0.315*	-0.140	
	(0.133)	(0.138)	(0.129)	(0.169)	
SPD dummy	-0.204	-0.204	-0.102	0.054	
	(0.158)	(0.156)	(0.129)	(0.151)	
UMP dummy	-0.309	-0.309	-0.439	-4.535*	
	(0.391)	(0.409)	(0.824)	(2.083)	
Constant	1.712***	1.712***	1.713***	1.551***	1.538***
	(0.115)	(0.108)	(0.117)	(0.142)	(0.017)
Party Level					-2.136***
					(0.248)
Year Level					-1.382**
					(0.485)
Multilevel Constant					-1.222***
					(0.139)
R^2	0.266	0.266	.	0.704	.
Root Mean Squared Error	0.418	0.418	0.301	0.278	.
AIC	66.877	66.877	.	-44.192	60.388
BIC	85.105	85.105	.	-21.913	66.464
Observations	56	56	56	56	56

Standard errors in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.