

Appendix A. Regression results emotion analysis

	<i>Dependent variable:</i>			
	polarity		arousal	
	NL (1)	DK (2)	NL (3)	DK (4)
Left-Right	0.00003 (0.005)	0.020** (0.008)	0.077** (0.037)	-0.004 (0.071)
Absolute Left-Right	-0.015*** (0.005)	0.007 (0.008)	-0.011 (0.038)	-0.029 (0.072)
Year	0.014** (0.005)	-0.018** (0.008)	-0.005 (0.038)	0.102 (0.070)
Seat share	0.010** (0.005)	0.032*** (0.008)	0.120*** (0.034)	-0.186** (0.073)
Party leader	-0.037*** (0.010)	-0.032* (0.018)	0.271*** (0.071)	0.032 (0.162)
In government	0.019* (0.010)	0.061*** (0.017)	-0.043 (0.072)	0.291* (0.153)
Women	-0.016 (0.014)	0.009 (0.021)	0.398*** (0.102)	0.150 (0.186)
Constant	0.335*** (0.009)	0.260*** (0.011)	6.851*** (0.063)	7.351*** (0.101)
Observations	660	518	660	518
R ²	0.068	0.104	0.057	0.024
Adjusted R ²	0.058	0.092	0.047	0.011
Residual Std. Error	0.111 (df = 652)	0.165 (df = 510)	0.795 (df = 652)	1.480 (df = 510)
F Statistic	6.813*** (df = 7; 652)	8.454*** (df = 7; 510)	5.676*** (df = 7; 652)	1.800* (df = 7; 510)

Note:

* ** *** p < 0.01

Appendix B. Application 2: topics in speeches

Our second application demonstrates how to track the salience of topics, over time and between parties. To identify the topics in the speeches we ran a Latent Dirichlet Allocation topic model. A topic model takes a term-document matrix¹ as input and produces (1) probabilities that words belong to specific topics and (2) probabilities that topics are mentioned in documents. A topic model essentially structures the correlations between words taking into account specific assumptions regarding the distributions of word frequencies. The researcher sets the number of topics to be identified and needs to substantively interpret the dimensions. Following a standard procedure, we set the number of topics to 35.² We subsequently identified topics by manual inspection of the words that loaded highly on these topics. Scholars have argued that the capacity and complexity of parties' agendas have increased (Green-Pedersen 2007; 2010), using a structural topic model (Roberts et al. XXX) would allow researchers to test whether this finding in manifesto's also holds up for congress speeches.

For clarity's sake, we focus here on the Dutch speeches and two topics: the EU and the economy. One of the 35 topics was easily identified as a Europe or EU topic.³ Four topics were identified as broadly covering different aspect of the economy. To calculate the probability that a speech covers the economy we added the probabilities of these four topics. Figure 2 panel A shows the 95% confidence interval of the probability that the European integration topic is mentioned in a given speech over time per 5-year interval (1945-2015). The prevalence of the topic dramatically increases over time, between 1945-1980 the probability is at best .01, in the 1990s and 2000s this respectively increases to .04 and .03. Panel B does the same but then for the economy. The economy is a more prevalent topic than the European integration. It peaks in the 1980s and 1990s and dramatically decreases again in the 2000s. Panel C shows the mean attention to European integration per party. Two pro-European integration parties (*GL* and *D66*) also speak the most about this

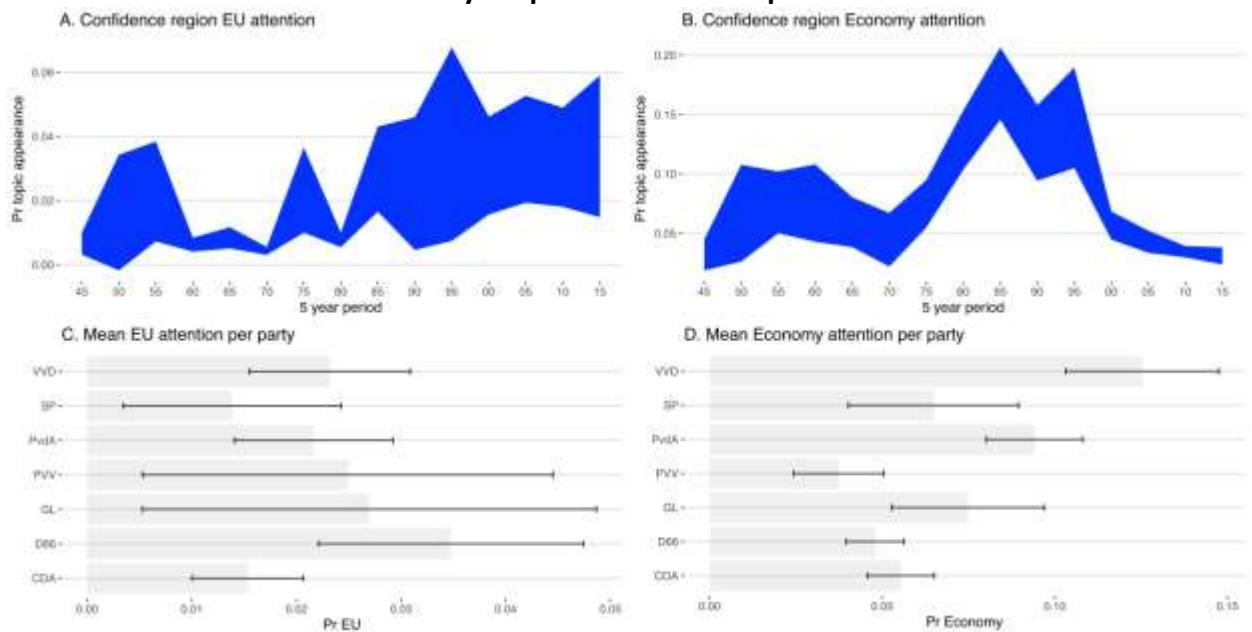
¹ This is a matrix with words as row, documents as column and cell entries represent the number of times a word is mentioned in a document. In our case the words were stemmed and the word list was not further preprocessed.

² We ran topic models with 5 to 80 topics, with incremental increases of 5 topics. For each topic model, we calculated the harmonic mean. This statistic indicates how much the expected word frequencies in the the model match the observed word frequencies. The goal is to find the best fit, i.e. the point after which adding new topics does not increase the match between observed and expected frequencies. This point was found at 35 topics.

³ High probabilities were given to words such as "europa", "europes", "europe", "parlement", "gemeenschap", "samenwerk".

topic, according to the model. The Socialist Party (*SP*) is surprisingly silent on the issue and the Christian Democrats (*CDA*) also score markedly lower than the other parties. Regarding the economy, two traditional economically left-wing (*PvdA*) and economically right-wing (*VVD*) parties have markedly more attention to the economy than the other parties. Parties such as the Freedom Party (*PVV*) and D66 that operate more on a cultural dimension also pay markedly less attention to the economy.

Figure 2. Attention to EU and the economy in speeches of Dutch parties



Note: panel A presents the probability that (part of) a speech is about the EU averaged over all speeches in 5 year time period. The confidence region spans the 95% confidence interval around the mean. Similarly, panel B presents the probability that (part of) a speech is about the economy. Panels C and D present the mean probabilities of the EU and Economy per party. The bar is the mean probability, and the lines are the 95% confidence intervals.