

Appendix A- Additional information on sample and coding of Bundestag sentences

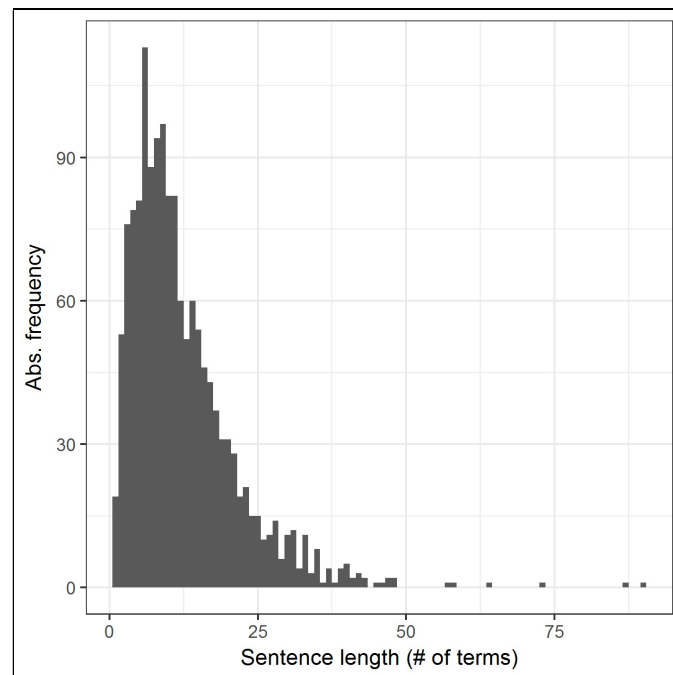


Figure A1: Length of sampled sentences

http://127.0.0.1:3827 Open in Browser Publish

Experiment zur Gewinnung der Grundstimmung in deutschsprachigen politischen Texten. Durchgang 1.

Kontakt: [Christian Rauh](#)

Details zu Forschungszielen und Ergebnissen sind nach Umfrageteilnahme vom Projektleiter verfügbar.

Die Abfragemaske basiert auf der Vorlage von Francis Smart.

[Survey-Tool](#)

Hits: 4

Willkommen!

Im Folgenden werden Dir 250 zufaellig ausgewaehte Textfragmente gezeigt.

Gib jeweils an, ob der Text fuer Dich eine neutrale, positive oder negative Grundstimmung transportiert.

Entscheide ganz spontan und ohne ueber die konkreten Textinhalte oder Kontexte nachzudenken. Es gibt kein richtig oder falsch, es geht nur um Deinen allerersten Eindruck.

Lade bitte die Ergebnisse am Ende dieser Umfrage herunter und schick sie mir per E-Mail (mit Name und Durchgang) zurueck.

Vielen Dank!

Weiter

http://127.0.0.1:3827 Open in Browser Publish

Experiment zur Gewinnung der Grundstimmung in deutschsprachigen politischen Texten. Durchgang 1.

Kontakt: [Christian Rauh](#)

Details zu Forschungszielen und Ergebnissen sind nach Umfrageteilnahme vom Projektleiter verfügbar.

Die Abfragemaske basiert auf der Vorlage von Francis Smart.

[Survey-Tool](#)

Hits: 4

4 / 250

halten sie eine bearbeitungszeit von 53 monaten seit dem beschluss des bundestages fuer ein raum- und funktionsprogramm des bundeskanzleramtes, das man normalerweise in drei monaten zustande bringt, fuer ein zeichen dafuer, wie zuegig sich die bundesregierung um die umsetzung des mehrheitsbeschlusses des deutschen bundestages bemueht ? .

Dein Eindruck:

☒ Neutral

☐ Positiv

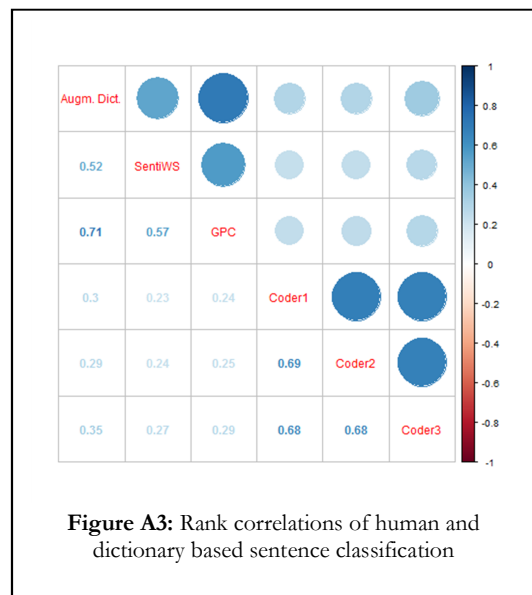
☐ Negativ

Weiter

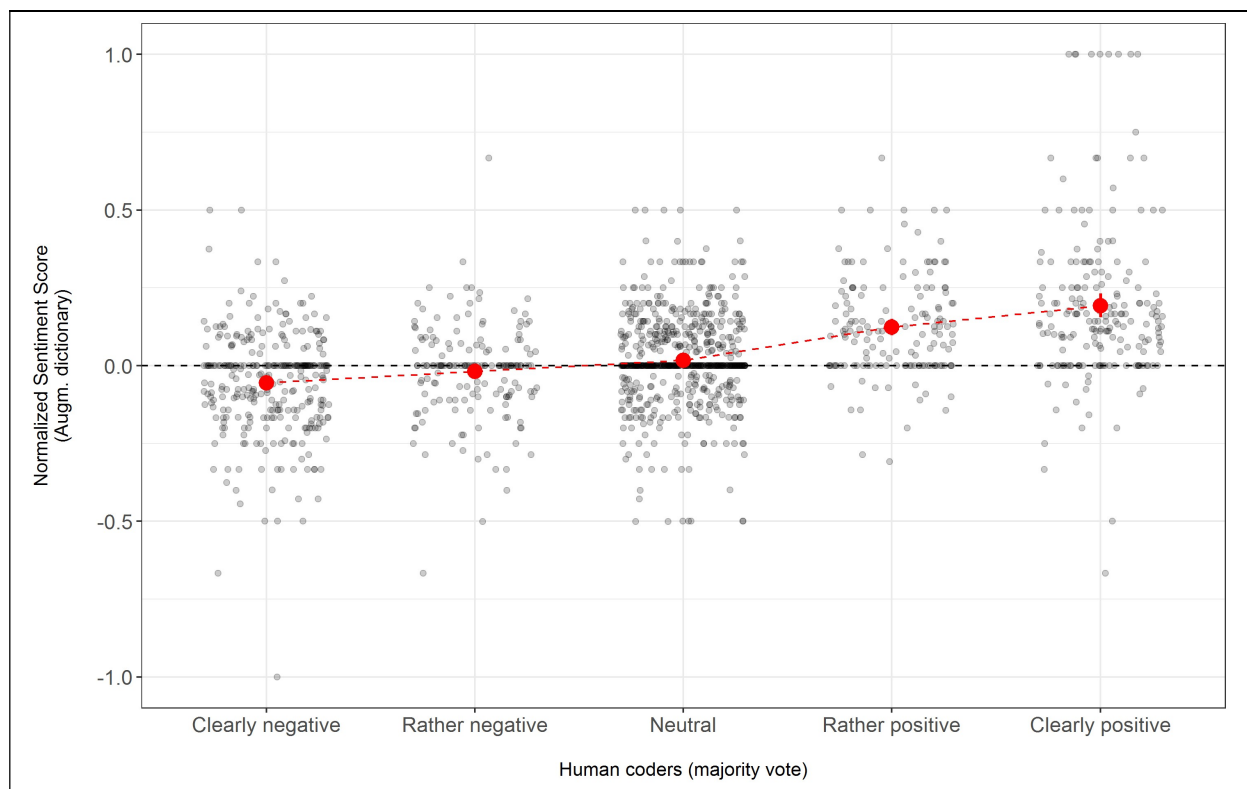
Figure A2: Screenshots of survey tool for human coders

Coders		Percent. agreement	Krippendorff α (ord. data)
Human coders	(3)	59.1	.675
Sent. dictionaries	(3)	77.2	.599
Humans & dictionaries	(6)	24.4	.404

Table A1: Inter-coder agreement
Note: 1.500 randomly sampled sentences from Bundestag speeches coded as ‘positive’, ‘neutral’, or ‘negative’.



Note: Despite their common roots also the sentiment dictionaries achieve no significantly better ‘inter-coder reliability’ than humans. The univariate distributions shown in Figure 1 in the main text together with the one-standard deviation cut-off-point for text classification suggest that particularly the *SentiWS* dictionary with its much lower spread disagrees frequently with the other two automated procedures. For applied research wishing to reduce automated scores to discrete ordinal scales this means that the robustness of the resulting inferences should be checked against different cut-off points.



Appendix B- Additional information on scoring Manifesto units

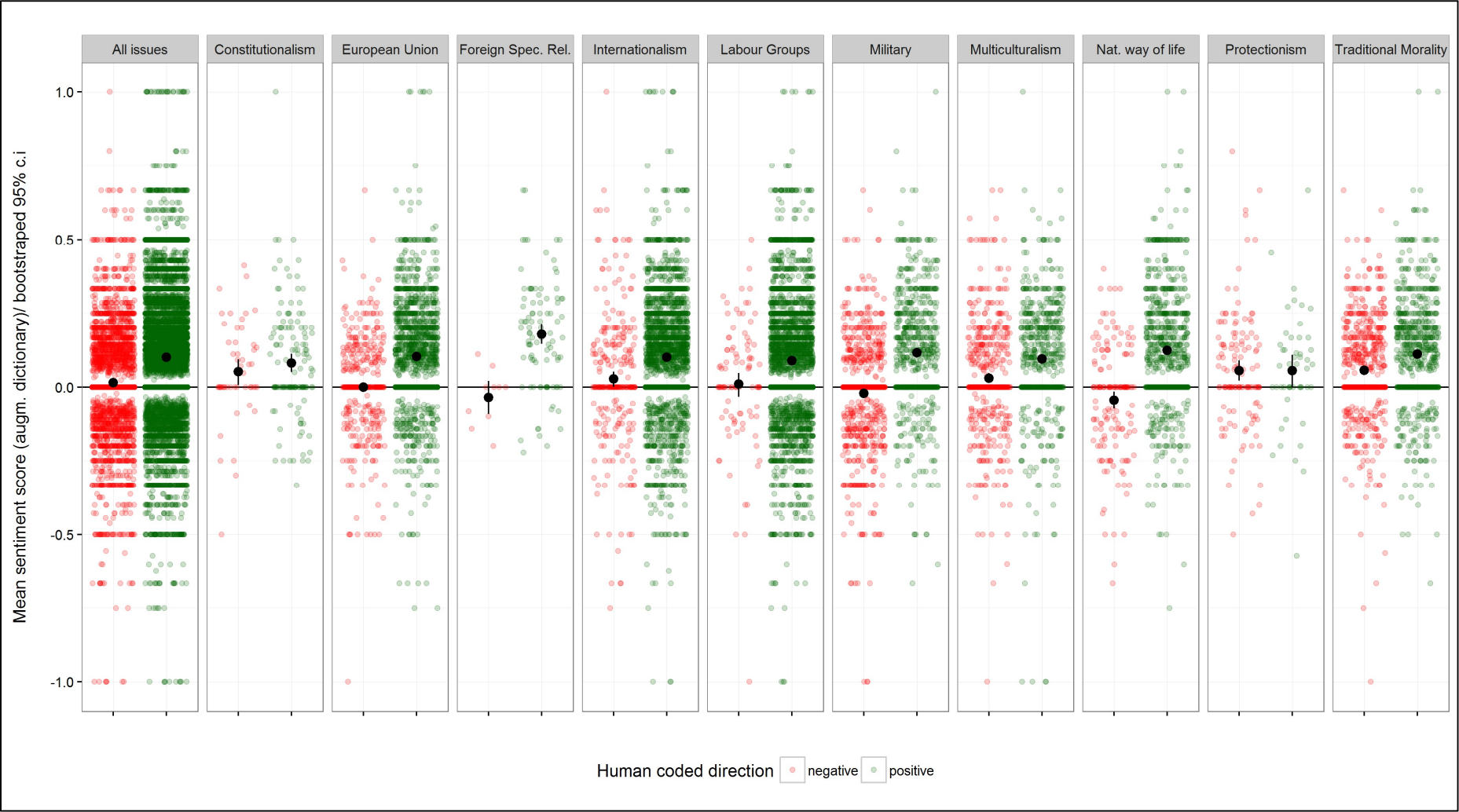


Figure B1: Full distribution of sentiment scores across human codes and Manifesto categories

Appendix C- Additional information on sampling/scoring ‘Causa Wulff’ media coverage

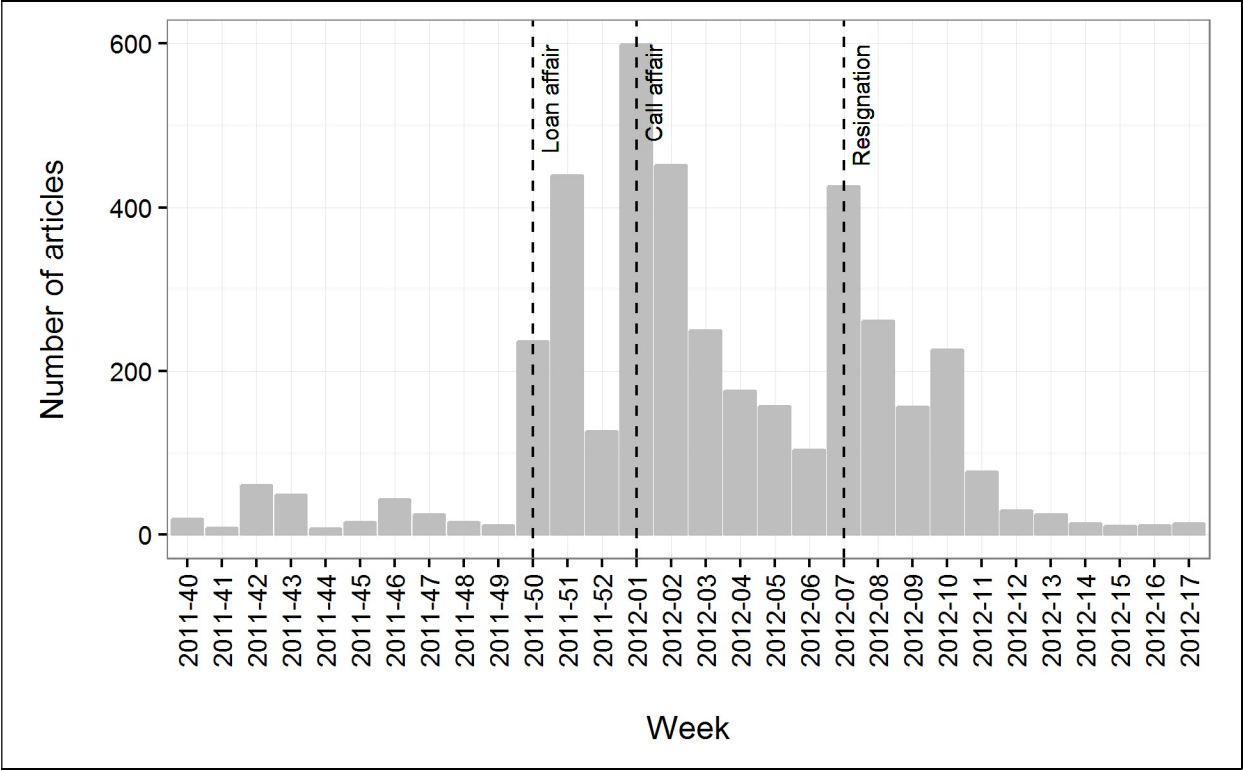


Figure C1: Temporal distribution of newspaper articles in ‘Causa Wulff’ sample

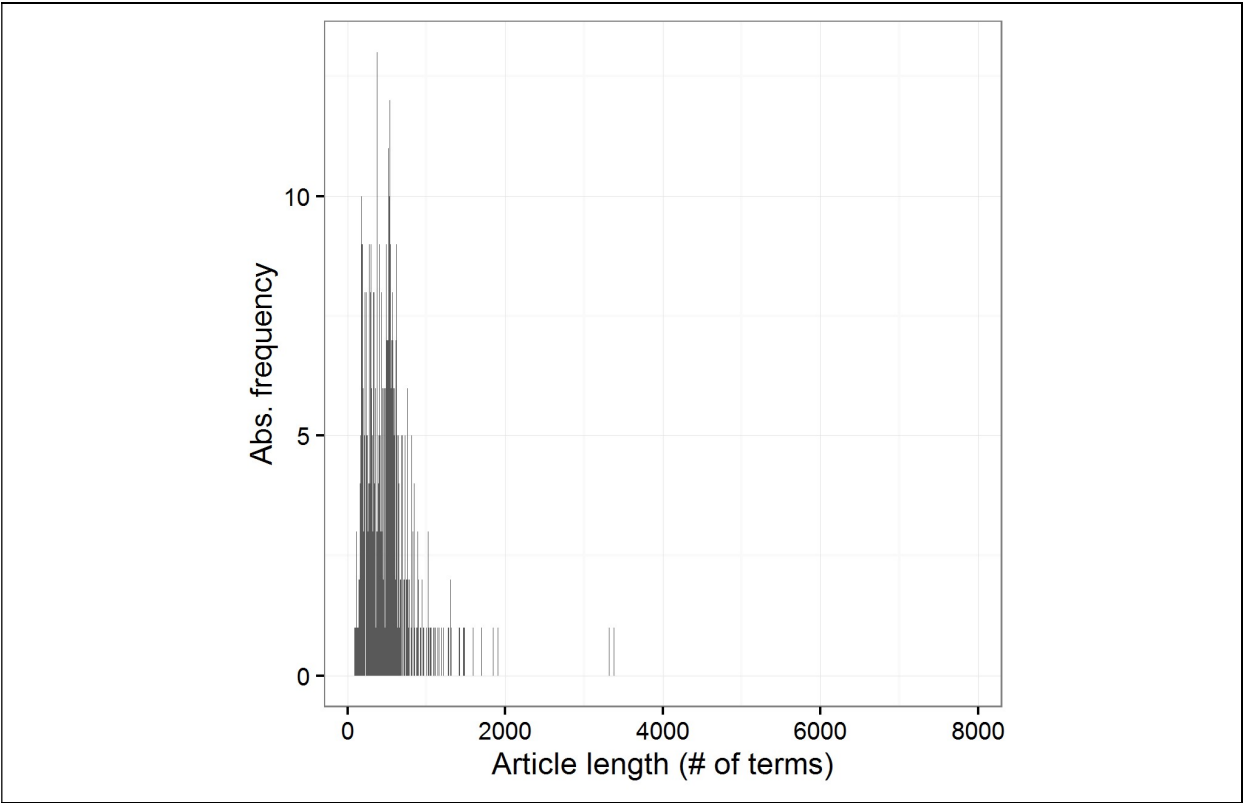


Figure C2: Histogram of article lengths in ‘Causa Wulff’ sample

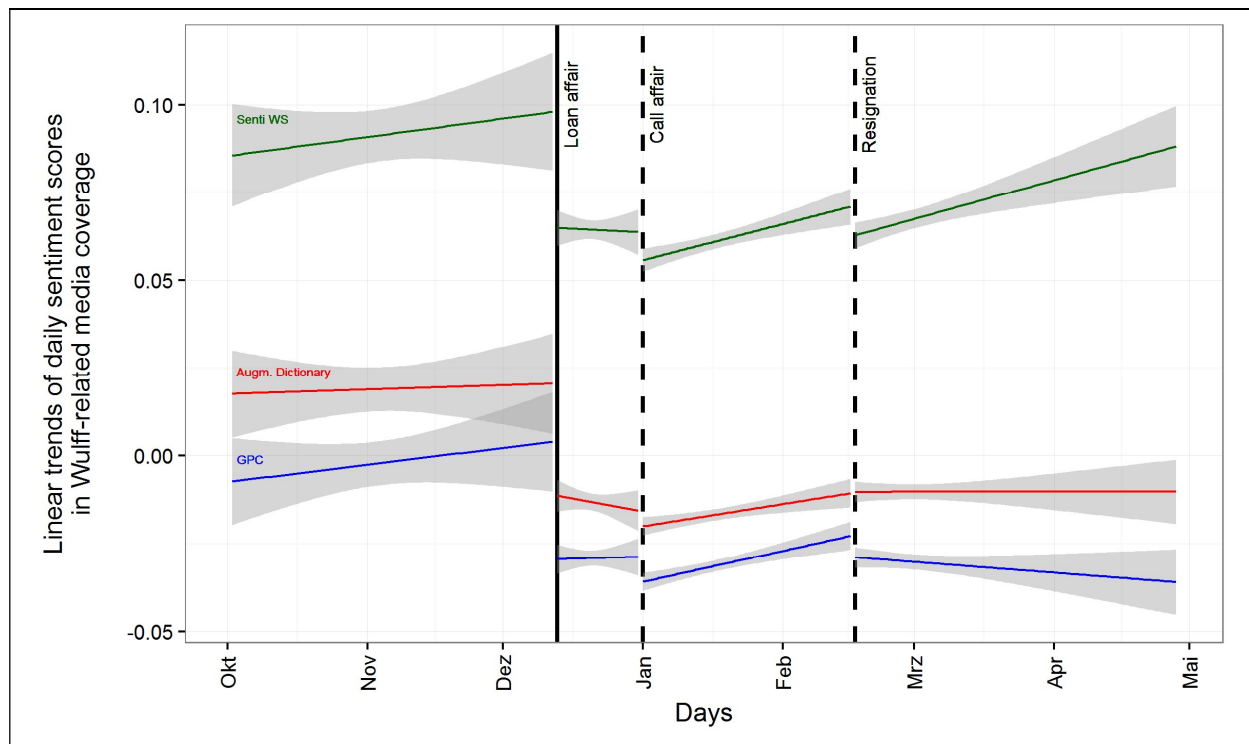


Figure C3: Comparing different dictionaries during the 'Causa Wulff'