The ties that bind.
Text similarities and the conditional diffusion of party policies.

Documentation

August 2020

Table of Contents

Introduction 1

Variable description 2

From source to dataset 3

References 4

Introduction

In the Dataverse we provide the following files to replicate the main analysis. All scripts were written under R versions 3.6.x.

File	Description
documentation.pdf	File description and codebook for dataset (this document)
dyadic_data_for_analyses.Rdata	Analysis dataset (R format)
regression_models_bjpols.R	R script to replicate the analysis and values mentioned in text
table02_forWord_regression_models.html	Output written by regression_models_bjpols.R (Table 2 in text)
figure1_vote.svg	Output written by regression_models_bjpols.R (Figure 1 in text)

Furthermore, we provide R Markdown/HTML files that form the base for the Supporting Information (SI), but go beyond. The following table shows the mapping from SI to Markdown.

Chapter in SI	Topic	Corresponding Markdown file
Appendix A	Making use of machine translation	suppl01_validation_report.html
Appendix B	Descriptive statistics	suppl02_descriptive_statistics.html
Appendix C	Regression models for Figure 2	suppl03_alt_model_specs.html
Appendix D	Alternative model specifications	
.	List of parties included	suppl04_party_list.html
Appendix E	Text similarity and left-right scores	suppl05_cosine_and_RILE.html
Appendix F	Comparing Jaccard and cosine similarity	suppl06_jaccard_similarity.html

Variable description

Originating from monadic data with party-elections as unit of observation, the analysis dataset has dyadic $party_{receiver}$-election ↔ $party_{sender}$-election as the unit of observation with the sender-election always preceding the receiver-election in time. A note on nomenclature, suffixes and prefixes:

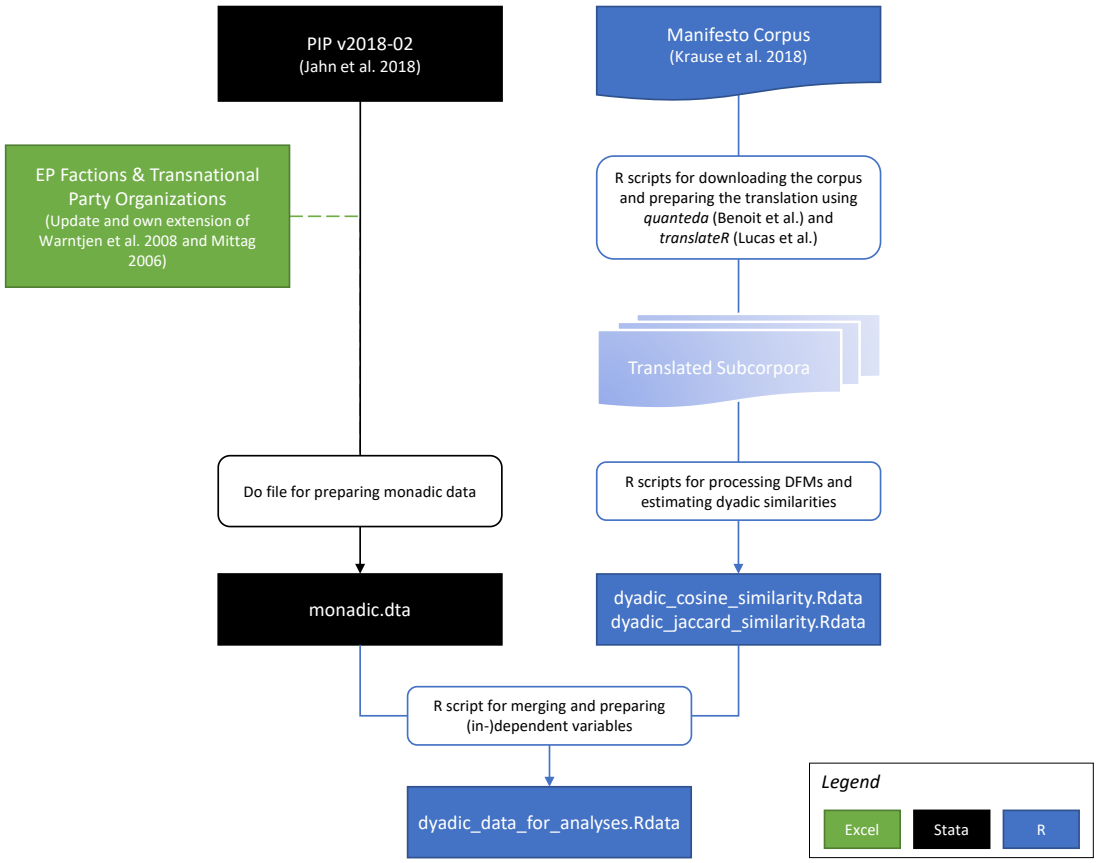
- Variables suffixed `.i` and `.j` refer to the receiving party *i* (R) and sending party *j* (S)
- t0 = focal election (e.g. no. 10)
- t-1 = past election (e.g. no. 9)
- t-2 = previous election (e.g. no. 8)
- l1 = lagged variable
- d1 = change between t0 and t-1 → "past change/move"
- d2 = change between t-1 and t-2 → "previous change/move" (cf. Adams et al. 2004)

Variable	Description
iso.i	R: Numeric country code (ISO 3166)
party.i	R: Party code according to Manifesto Project
elecid.i	R: Unique code for each election (i.e. "Iso.ElecNo")
decade.i	R: Decade of election
year.i	R: Year of election
famofnation.i	R: Family of nations
iso.j	S: Numeric country code (ISO 3166)
party.j	S: Party code according to Manifesto Project
elecid.j	S: Unique code for each election (i.e. "Iso.ElecNo")
decade.j	S: Decade of election
year.j	S: Year of election
famofnation.j	S: Family of nations
d1_landslide.i	R: Relative gain/loss as p% of *i*'s past vote share
d1_gain.i	R: Dummy: 1 if *i* gained votes
d1_vote.i	R: Δ of vote share (t0 and t-1)
p310.i	R: Vote share at t0 (level)
d2_landslide.i	R: Relative gain/loss as p% of *i*'s previous vote share
d2_gain.i	R: Dummy: 1 if *i* gained votes
d2_vote.i	R: Δ of vote share (t-1 and t-2)
l1_voteshare.i	R: Vote share at t-1 (level)
d1_landslide.j	S: Relative gain/loss as p% of *j*'s past vote share
d1_gain.j	S: Dummy: 1 if *j* gained votes
d1_vote.j	S: Δ of vote share (t0 and t-1)
p310.j	S: Vote share at t0 (level)
d2_landslide.j	S: Relative gain/loss as p% of *j*'s previous vote share
d2_gain.j	S: Dummy: 1 if *j* gained votes

Variable	Description
d2_vote.j	S: Δ of vote share (t-1 and t-2)
l1_voteshare.j	S: Vote share at t-1 (level)
re_use	Dummy: 1 if *i* is linked to itself at the past election (i.e. $i_t = i_{t-1}$)
competitor	Dummy: 1 if *j* was a domestic competitor at the past election
same_famofnat	Dummy: 1 if *i* and *j* are in the same "family of nations"
same_gov	Dummy: 1 if *i* and *j* are both government parties
same_epfaction	Dummy: 1 if *i* and *j* are members of the same EP faction
same_ipo	Dummy: 1 if *i* and *j* are members of the same TPO
neighbor	Dummy: 1 if *i* and *j*'s countries share a border
sender_gov	Dummy: 1 if *j* is in government but not *i*
cosine_similarity	Cosine similarity of both manifestos
jaccard_similarity	Jaccard similarity of both manifestos
bu01o.i	R: RILE score
bu01o.j	S: RILE score
abs_diff_RILE	Absolute difference of RILE scores

From source to dataset

The following flow chart shows how to construct the analysis dataset from source. R scripts and do-files are available upon request from the authors.



References

Adams J *et al.* (2004). Understanding Change and Stability in Party Ideologies: Do Parties Respond to Public Opinion or to Past Election Results? *British Journal of Political Science* **34** (4), 589–610.

Benoit K *et al.* (2018). quanteda: An R package for the quantitative analysis of textual data. *The Journal of Open Source Software* **3** (30), 774.

Jahn D *et al.* 2018. PIP – Parties, Institutions & Preferences: PIP Collection [Version 2018-02]. University of Greifswald.

Krause W *et al.* 2018. Manifesto Corpus: Version: 2018-2. Wissenschaftszentrum Berlin (WZB).

Lucas C *et al.* (2015). Computer-Assisted Text Analysis for Comparative Politics. *Political Analysis* **23** (2), 254–277.

Mittag, J ed. (2006). *Politische Parteien und europäische Integration: Entwicklung und Perspektiven transnationaler Parteienkooperation in Europa.* Essen: Klartext.

Warntjen A, Hix S and Crombez C (2008). The Party Political Make-Up of EU Legislative Bodies. *Journal of European Public Policy* **15** (8), 1243–1253.