Data and Discourse around Europe

Wolf J. Schünemann

Darmstadt, 20 Feb 2020

Plan

1   Data and computational methods in political science

2   European politics and the second order question

3   Election campaigns in the transnational campaign arena

4   Comparative EU referendum research

5   Experimental application of (Structural) Topic Modelling

6   Conclusions

1   Data and computational methods in political science

"This is a world where massive amounts of data and applied mathematics replace every other tool that might be brought to bear. Out with every theory of human behavior, from linguistics to sociology. Forget taxonomy, ontology, and psychology. Who knows why people do what they do? The point is they do it, and we can track and measure it with unprecedented fidelity. With enough data, the numbers speak for themselves."

"There's no reason to cling to our old ways. It's time to ask: What can science learn from Google?"

Anderson, Chris: The End of Theory: The Data Deluge Makes the Scientific Method Obsolete. URL: https://www.wired.com/2008/06/pb-theory/(21.03.2019).

1   Data and computational methods in political science

Opportunities

- Social Sciences in general and political science in particular interested in social relations
- Datafication and CSS provide social scientists with the means to study large amounts of 'natural' data using non-intrusive methods
- Interest in relations between actors → power/influence → network analysis using meta-data (Lazer & Wojcik 2018; Ruths & Pfeffer 2014); Discourse *Network* Analysis (Leifeld 2018)
- Interest in discourse → power/knowledge relations → text-as-data (Diaz-Bone & Widmer 2018; Glavaš et al.)

Diaz-Bone, Rainer; Widmer, Jennifer (2018): Zum Stand der Diskursforschung - Anmerkungen und Befunde zu Institutionalisierung, Problemen und Struktur eines transdisziplinären Feldes. In: Zeitschrift für Diskursforschung 2018 (2), S. 146–161. Glavaš, G., Nanni, F., & Ponzetto, S. P. (2019). Computational Analysis of Political Texts: Bridging Research Efforts Across Communities. In P. (H.) Nakov (Ed.), Proceedings of the 57th Annual Meeting of the ACL (pp. 18–23). Lazer, D., & Wojcik, S. (2018). Political Networks and Computational Social Science. In J. N. Victor, A. H. Montgomery, & M. Lubell (Eds.), The Oxford handbook of political networks (pp. 115–130). New York, NY: Oxford University Press. Leifeld, P. (2018). Discourse Network Analysis. In J. N. Victor, A. H. Montgomery, & M. Lubell (Eds.), The Oxford handbook of political networks (pp. 301–325). New York, NY: Oxford University Press. Ruths, D., & Pfeffer, J. (2014). Social sciences. Social media for large studies of behavior. *Science (New York, N.Y.)*, 346(6213), 1063–1064.

1 Data and computational methods in political science

Risks / Challenges

- Production of artefacts and spurious correlations (see election forecasts with Google data Tumasjan et al. 2011)
- Uninformed interpretation of computationally produced output without sufficient theoretical underpinnings or knowledge of the research object (e.g. network graphs, Topic Modelling)
- Unexplainable or Black-box-AI
- "Mathwashing" of subjective findings (Eubanks 2019)
- Purely data-driven research only little acceptance in more traditional branches of political science
- Risk of becoming a niche with most innovation coming from computer science and computer linguists

→ ideal: answer meaningful questions in explainable ways

Eubanks, V. (2019). Automating inequality: How high tools profile, police, and punish the poor (First Picador edition). New York: Picador; St. Martin's Press. Tumasjan, A., Sprenger, T. O., Sandner, P. G., & Welpe, I. M. (2011). Election Forecasts With Twitter. Social Science Computer Review, 29(4), 402–418.

Illustration „Die Automate", E.T.A. Hoffmann, URL: http://deutscheliteratur.net/e-t-a-hoffmann-der-sandmann-1816/

2   European politics and the second-order question

European Parliament elections

- Second-order hypothesis in EU studies means that European politics plays a secondary part (at best) in comparison to national politics for all relevant actors: the electorate, politicians, media etc.
- At least two dimensions: transnational activity and EU-related issue salience or even politicization
- Derived from empirical research on EU parliament elections (Reif & Schmitt 1980), repeatedly and convincingly tested for successive elections (Hix & Marsh, 2011; Marsh, 1998; Schmitt, 2005; Träger, 2015)
- EU issues have become more politicized (Kriesi & Grande 2016)
- Transnational personalization through the Spitzenkandidaten system (Schmitt et al. 2015)
- Social media might enhance transnational public sphere (Bossetta et al., 2017; Hänska & Bauchowitz, 2019)

Bossetta, M., Dutceac Segesten, A., & Trenz, H.-J. (2017). Engaging with European Politics Through Twitter and Facebook: Participation Beyond the National? In M. Barisione & A. Michailidou (Eds.), Social Media and European Politics: Rethinking Power and Legitimacy in the Digital Era (pp. 53-76). London: Palgrave Macmillan UK. Hänska, M., & Bauchowitz, S. (2019). Can social media facilitate a European public sphere? Transnational communication and the Europeanization of Twitter during the Eurozone crisis. Social Media + Society, 5(3). Hix, S., & Marsh, M. (2011). Second-order effects plus pan-European political swings: An analysis of European Parliament elections across time. Electoral Studies, 30(1), 4–15. Kriesi, H., & Grande, E. (2016). The euro crisis: a boost to the politicisation of European integration? In S. Hutter, E. Grande, & H. Kriesi (Eds.), Politicising Europe: Integration and mass politics (pp. 240–276). Cambridge, United Kingdom: Cambridge University Press. Marsh, M. (1998). Testing the Second-Order Election Model after four European Elections. British Journal of Political Science, 28(04), 591–607. Reif, K. H., & Schmitt, H. (1980). Nine Second-Order National Elections: A Conceptual Framework for the Analysis of European Election Results. European Journal of Political Research, 8(1), 3–44. Schmitt, H. (2005). The European Parliament Elections of June 2004: Still Second-Order? West European Politics, 28(3), 650–679. Schmitt, H., Hobolt, S., & Popa, S. A. (2015). Does personalization increase turnout? Spitzenkandidaten in the 2014 European Parliament elections. European Union Politics, 16(3), 347-368. Träger, H. (2015). Die Europawahl 2014 als second-order election: Ein Blick in alle 28 EU-Staaten. In N. Switek (Ed.), Die Europawahl 2014: Spitzenkandidaten, Protestparteien, Nichtwähler (pp. 33–44). Wiesbaden: Springer VS.

EU referendums

- Issue-voting against second order also the most prominent dispute in EU referendum research
- Very resounding in the political/normative debate on EU referendums, discursive strategy in ex-post rationalization of 'failed' EU referendums (see also Brexit)
- Transferred to EU referendums most prominently by Franklin et al. (so-called "Franklin thesis", Eijk, C. van der & Franklin, 1996; Franklin, Marsh, & McLaren, 1994); tested (or argued for) for the referendums in the constitutional process (Le Gall, 2005; Moravcsik, 2005; Nijeboer, 2005)
- Disputed by many scholars finding support for issue-orientation for different cases (Beach, 2009; Siune, Svensson, & Tongaard, 1994; Svensson, 2002)
- Classification between second-order and issue-voting often over-simplified (notable exception Glencross & Trechsel, 2011)

Eijk, C. van der, & Franklin, M. (Eds.) (1996). *Choosing Europe. The European electorate and national politics in the face of union.* Ann Arbor: Univ. of Michigan Press. Franklin, M., Marsh, M., & McLaren, L. (1994). Uncorking the Bottle: Popular Opposition to European Unification in the Wake of Maastricht. Journal of Common Market Studies, 32(4), 455–472. Le Gall, G. (2005). Le 29 mai entre mythe et réalités. Revue Politique Et Parlementaire. (1036), 2–23. Moravcsik, A. (2005). Europe without illusions: a category error. Prospect. (112), 1–9. Nijeboer, A. (2005). The Dutch Referendum. European Constitutional Law Review. (1), 393–405. Nijeboer, A. (2005). The Dutch Referendum. European Constitutional Law Review. (1), 393–405.

Source: The Guardian

Remain **48%** 16,141,241

Leave **52%** 17,410,742

0 of 382 local authorities still to declare

72.16% turnout

Beach, D. (2009). The costs of no - a two-dimensional issue-voting model of voter behavior in EU referendums. In EUSA Eleventh Biennial Conference, Los Angeles. Svensson, P. (2002). Five Danish referendums on the European Community and European Union: A critical assessment of the Franklin thesis. European Journal of Political Research, 41(6), 733–750. Siune, K., Svensson, P., & Tongaard, O. (1994). The EU: The Danes said 'No' in 1992, but 'Yes' in 1993: How and Why? Electoral Studies, 13(2), 107–116. Svensson, P. (2002). Five Danish referendums on the European Community and European Union: A critical assessment of the Franklin thesis. European Journal of Political Research, 41(6), 733–750.

3　　Twitter communication in the transnational campaign arena

Research object and questions

- "Campaigning in the transnational arena. Candidates' transnational linkages on Twitter during the 2019 European Parliament elections" (together with Caterina Froio and Sebastian Stier)
- RQ1 To what extent is campaign communication by EP candidates transnational?
- RQ2 Under what conditions do EP candidates engage in transnational campaign activities?

Why Twitter (again)?
- Twitter as a unique data source (not only uniquely accessible)
- More data, more fine-grained
- Unmediated, less curated political **elite** communication
- Best available measurement of transnational linkages between political actors

3   Twitter communication in the transnational campaign arena

Dataset and method

Dataset
- 7,917 EP candidates from all 28 EU member states, of which
- N = 3,852 had a Twitter account (data collected with
- Euromanifesto Study)
- 3,070 EP candidates sent 566,948 tweets from 23 April – 30 May 2019
- All tweets by EP candidates bought from Twitter

Method
- Network analytical design: EP candidates as nodes, @-mentions and retweets from Tweets as nodes
- 40,018 @-mention dyads and 38,072 retweet dyads within our set of candidates
- Statistical measures of effects

3   Twitter communication in the transnational campaign arena
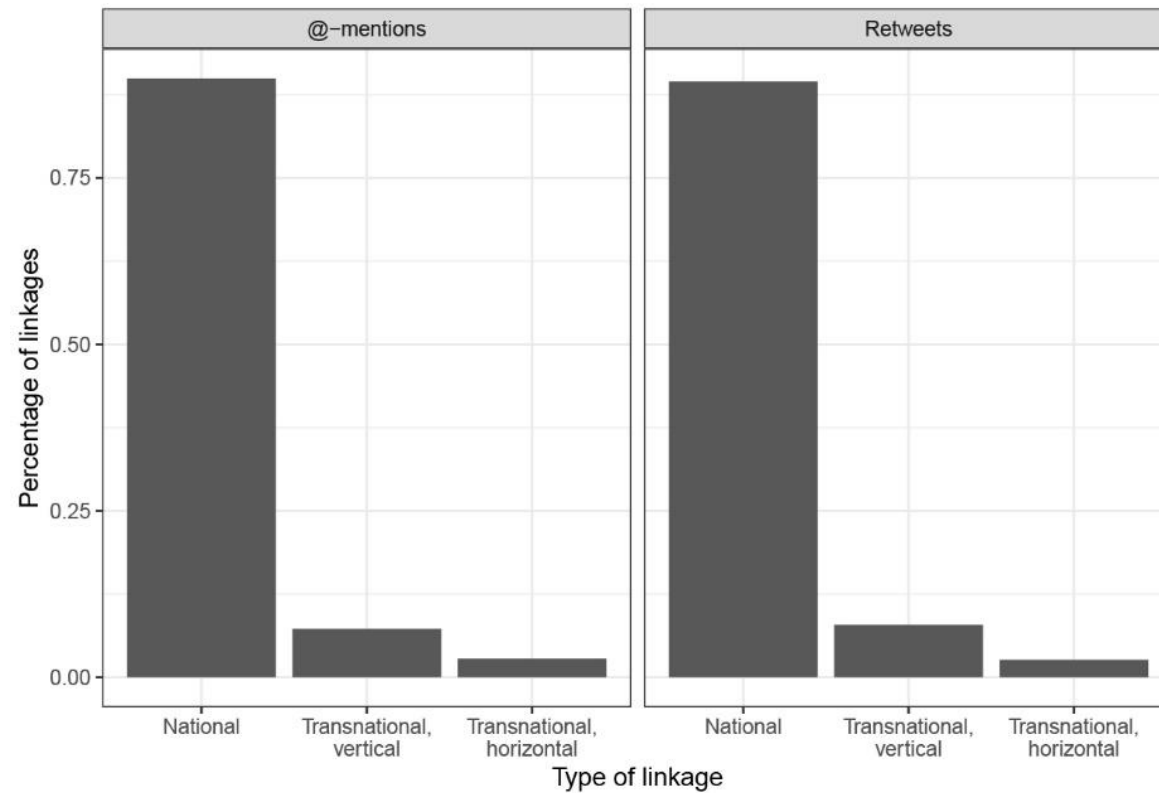## Theoretical conceptions

Dependent variables
- National: @-mentions or retweets of an EP candidate from the same country
- Transnational, horizontal: @-mentions or retweets of an EP candidate from another country
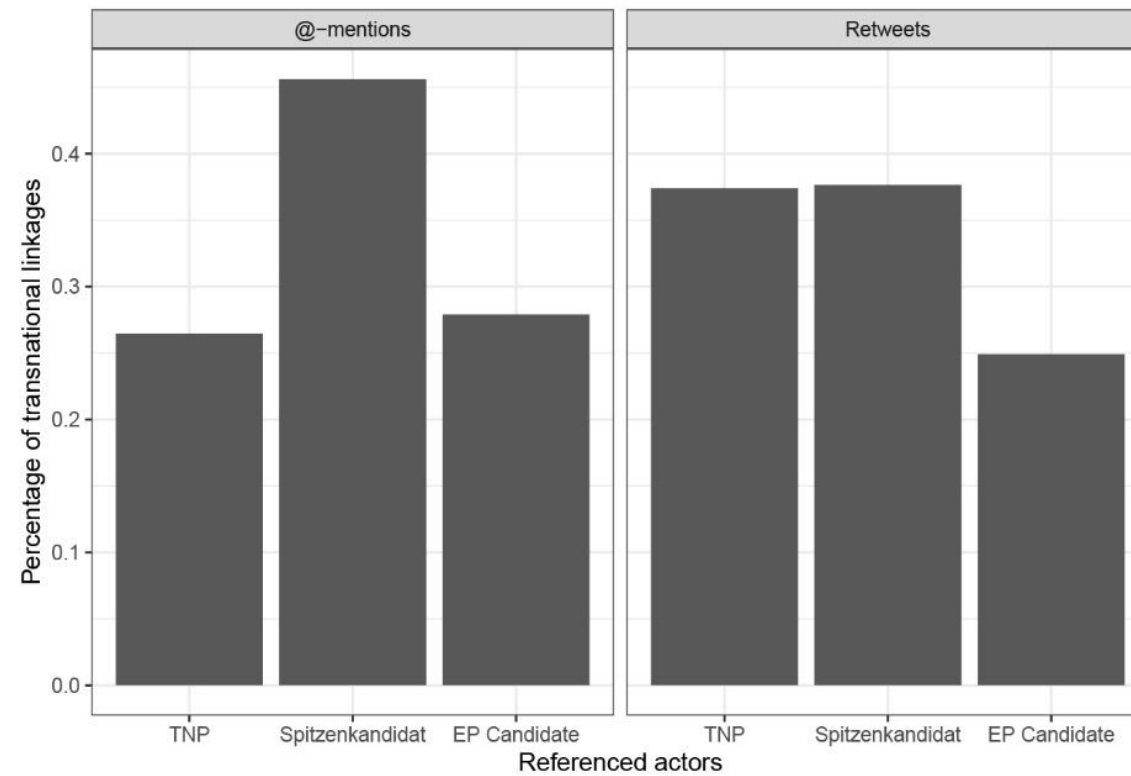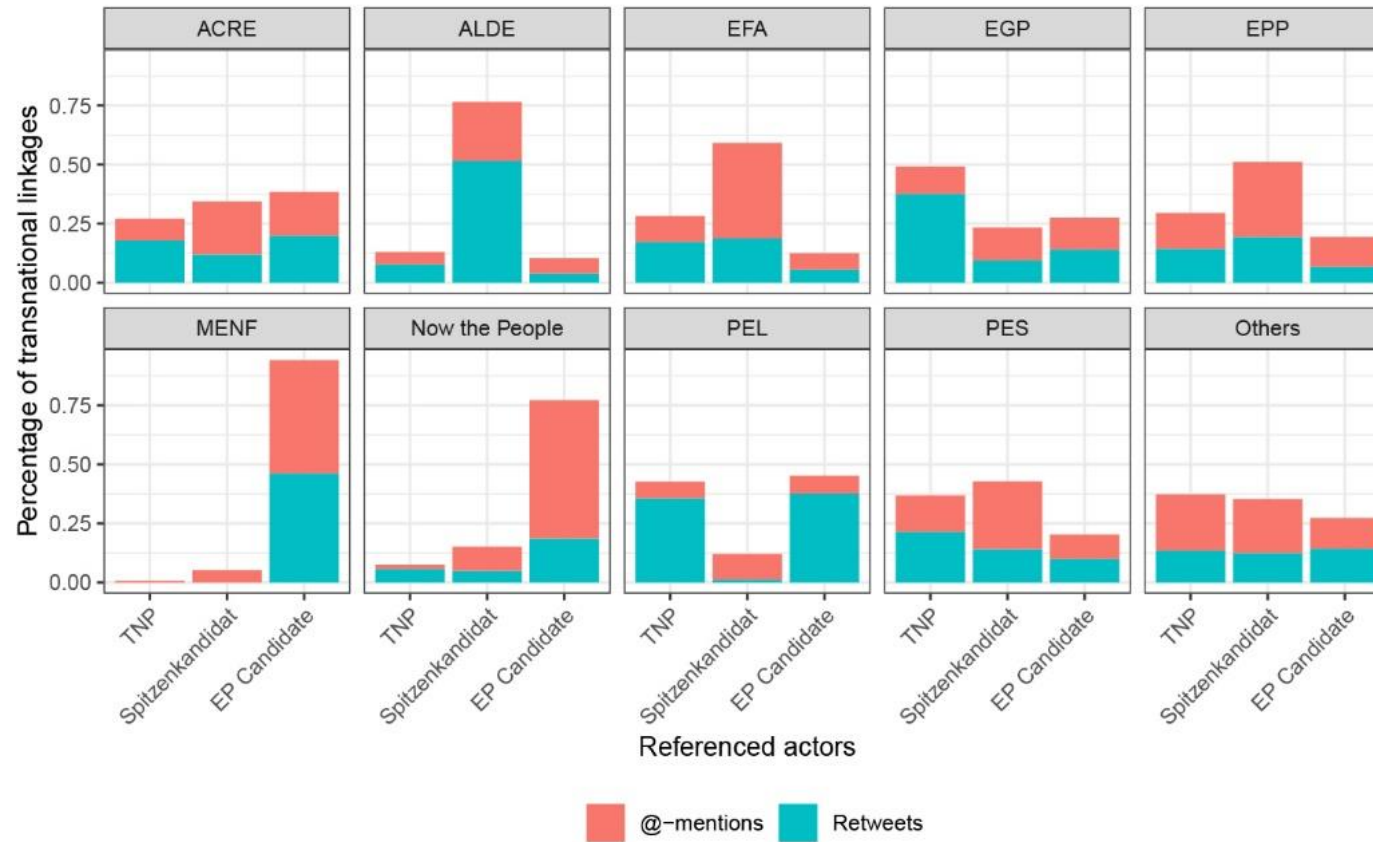- Transnational, vertical: @-mentions or retweets of EU TNPs or Spitzenkandidaten
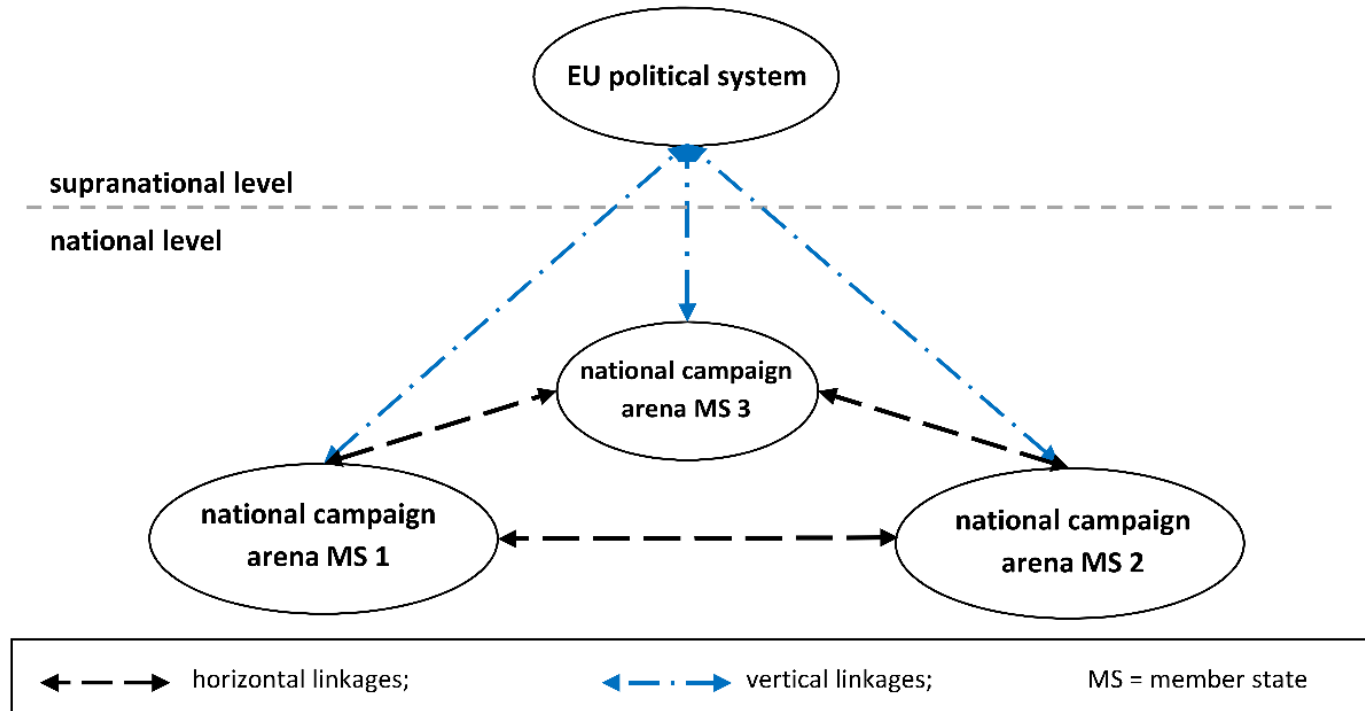
Results

3 Twitter communication in the transnational campaign arena
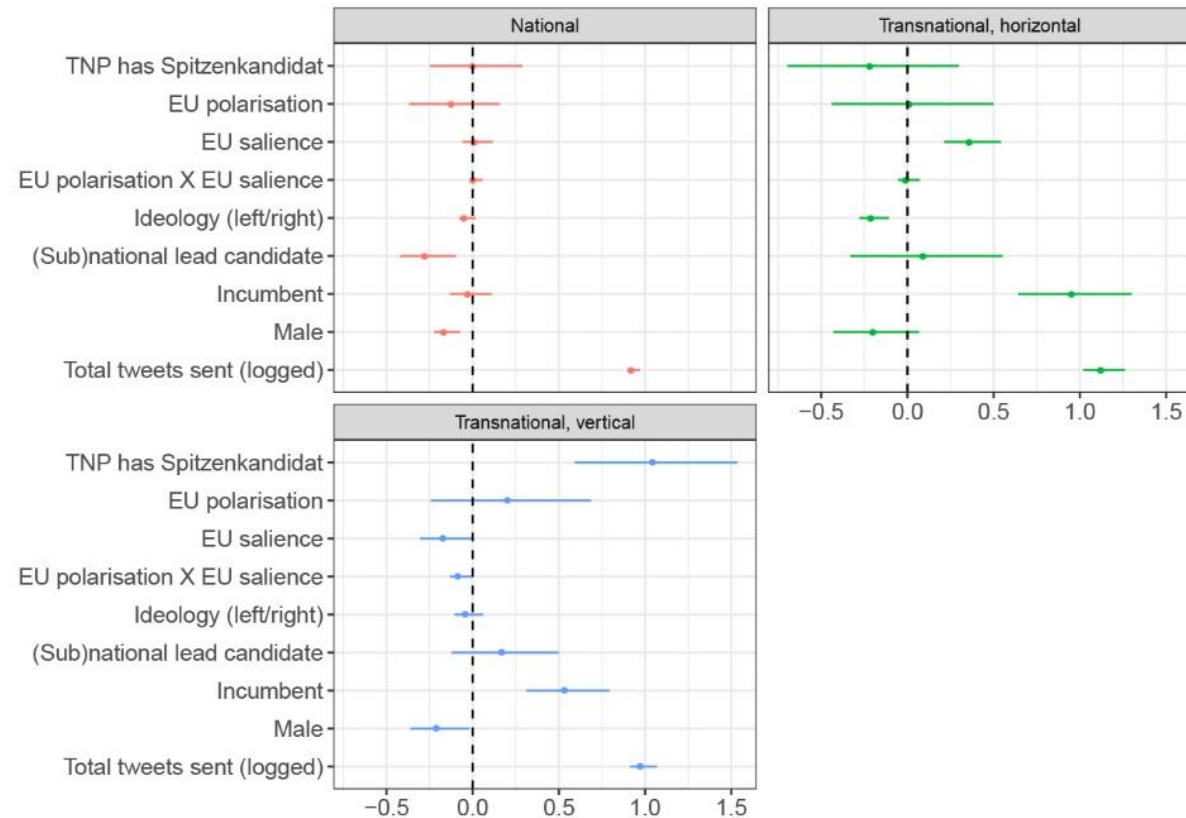
Results

3 Twitter communication in the transnational campaign arena

Results

3 Twitter communication in the transnational campaign arena

Results

Negative binomial regressions; candidates nested in parties; country dummies included.

3   Twitter communication in the transnational campaign arena

Preliminary conclusions

- EP campaigns remain mostly national
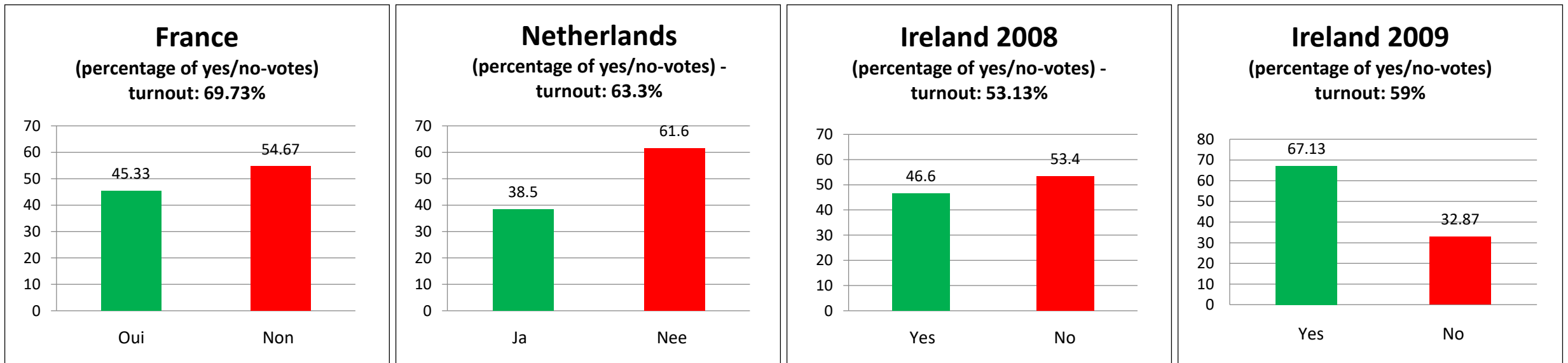- **support for second-order hypothesis for EP elections and transnational linkages**

BUT:

- Some telling variation among parties observable with signifikant effect of the Spitzenkandidaten for parties that nominated one (+ vertical transnational linkages)
- Comparison is needed beyond the election campaigns
- What about discourse analysis?

Comparative EU referendum research
Cases

France
(percentage of yes/no-votes)
turnout: 69.73%

45.33 — Oui
54.67 — Non

Netherlands
(percentage of yes/no-votes) -
turnout: 63.3%

38.5 — Ja
61.6 — Nee

Ireland 2008
(percentage of yes/no-votes) -
turnout: 53.13%

46.6 — Yes
53.4 — No

Ireland 2009
(percentage of yes/no-votes)
turnout: 59%

67.13 — Yes
32.87 — No

Sources (data):
1.   http://www.interieur.gouv.fr/sections/a_votre_service/resultats-elections/rf2005/000/000.html (16.7.2008)
2.   http://www.referendumplatform.nl/share/files/13_334315/gehoudenreferenda.pdf (s.o.)
3.   http://electionsireland.org/results/referendum/refresult.cfm?ref=2008R (s.o.)
4.   http://www.referendum.ie/current/index.asp?ballotid=79 (17.11.2009)

Comparative EU referendum research
Modified hypotheses

H1: EU treaty referendum campaigns are narrowly oriented towards the issue at hand, thus the legal provisions enshrined in the respective reform treaty.

H2: EU treaty referendum campaigns do not concentrate on the provisions enshrined in the respective reform treaty but include broader aspects of European integration.

H3: EU treaty referendum campaigns are second-order as they put emphasis on the performance of national political actors, especially government actors at cost of more issue-oriented arguments.

H4: Yes campaigns in EU treaty referendums show a higher degree of issue-orientation (treaty-based or more generally oriented towards EU integration) than no campaigns.

4　Comparative EU referendum research

Multiple methods and dataset

- Qualitative discourse analysis following the Sociology of Knowledge Approach to Discourse (Keller 2008, 2011, 2013; Schünemann 2014, 2017, 2018)

- Topic modelling: automated machine-learning approach, LDA algorithm (Blei 2012; Brett 2012), R packages tm & topicmodels

Dataset

- Campaign material (here: texts) of parties, interest groups and ad-hoc-campaign organisations produced during the campaign periods

	F 2005 Oui	F 2005 Non	NL 2005 Ja	NL 2005 Nee	IRL 2008 Yes	IRL 2008 No	IRL 2009 Yes	IRL 2009 No
texts	220	448	114	185	269	239	277	255
words	276,644	646,944	72,505	152,114	239.054	208,398	186,180	195,172

Comparative EU referendum research

Conception of the argument in political discourse
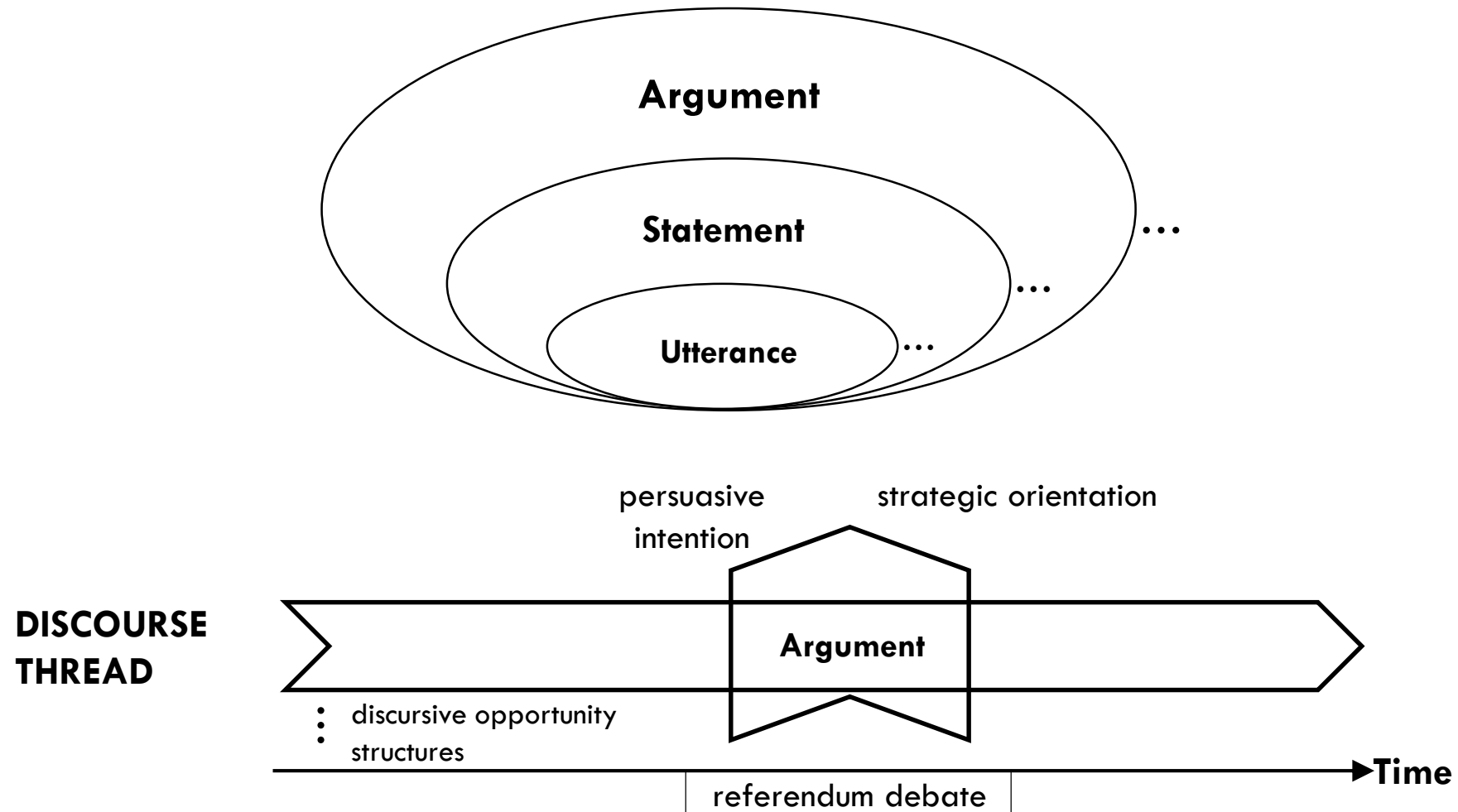
Comparative EU referendum research

Conception of the argument in political discourse

Yes-/No-discourse

Vote yes/no to the treaty,…
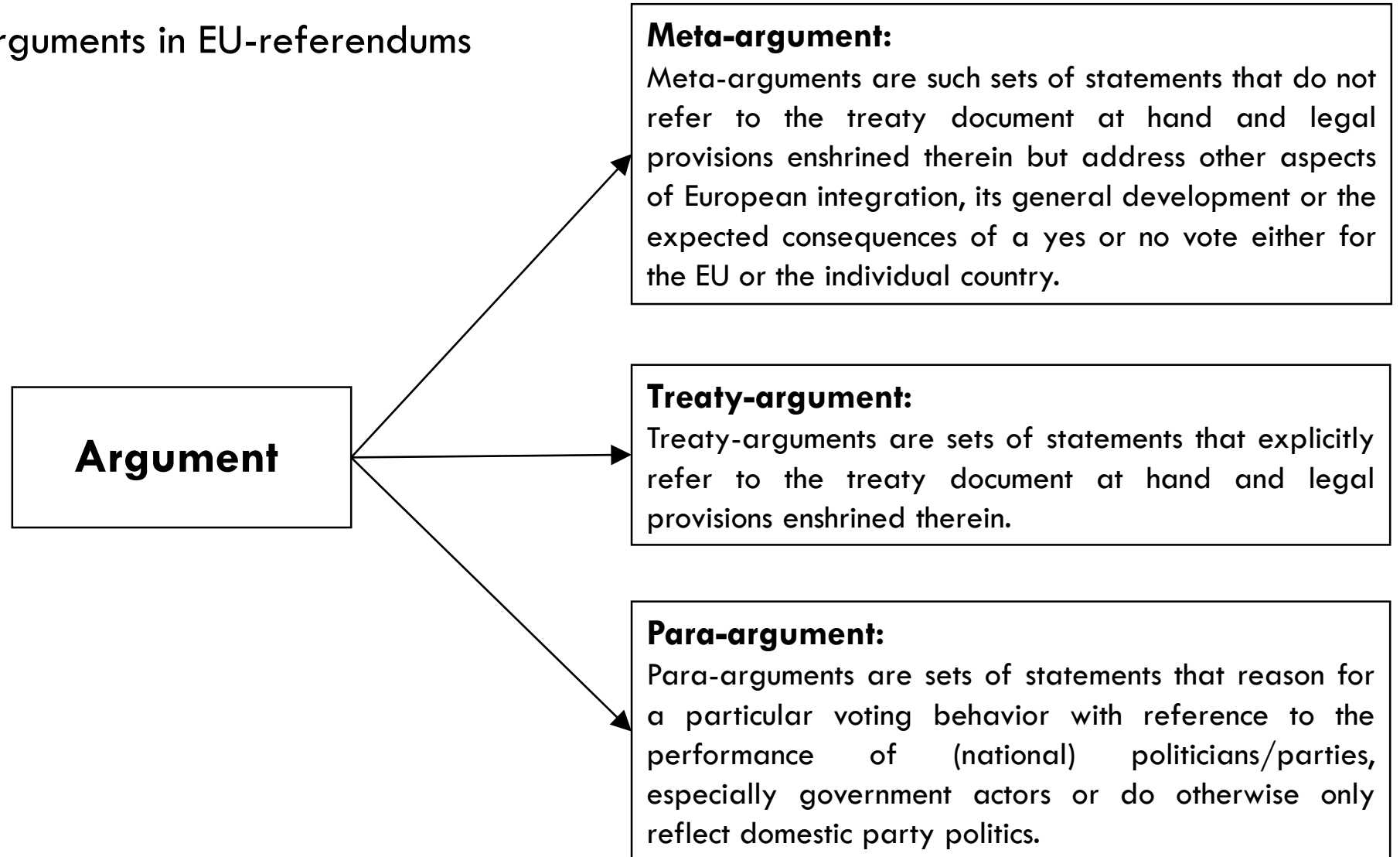
Why?

argument

…because…

Why?

sub-argument

…because…

Yes-discourse		No-discourse	
Argument (Ex.)	Sub-arguments	Argument (Ex.)	Sub-arguments
Democracy Argument: DemA: *The treaty makes the EU political process more democratic.*	ParlA: *The European Parliament will get more powers (s. articles x, y and z).*	Militarization Critique: MilC: *The treaty leads to a further militarization of the EU.*	SpeC: *The spectrum for EU missions will be extended (art. x).*
	SubsA: *National parliaments will get a stronger role in EU decision-making (art. x).*		DefC: *A European Defense Agency is established (art. x).*
	BIA: *A Citizen's initiative is to be introduced (art. x).*		ArmC: *Member states are obliged to increase their military budgets (art. y).*
	TransA: *The decision-making on EU-level will become more transparent (art. x).*		PermaC: The instrument of *Permanent Structured Co-operations is to be introduced (art. x).*

Types of arguments in EU-referendums

Argument

Meta-argument:
Meta-arguments are such sets of statements that do not refer to the treaty document at hand and legal provisions enshrined therein but address other aspects of European integration, its general development or the expected consequences of a yes or no vote either for the EU or the individual country.

Treaty-argument:
Treaty-arguments are sets of statements that explicitly refer to the treaty document at hand and legal provisions enshrined therein.

Para-argument:
Para-arguments are sets of statements that reason for a particular voting behavior with reference to the performance of (national) politicians/parties, especially government actors or do otherwise only reflect domestic party politics.

Types of arguments in EU-referendums – classification

Table 1: → Treaty references – yes camp France¶

Table 2: → Treaty references – no camp France¶

Argument	aspect	treaty provision / article

Social Policy Argument	¤	[treaty argument]¤
	social policy objectives¤	Union's objectives (art I-3); horizontal social clause (*clause sociale transversale*, art III-117)¤
	workers' rights¤	Right to engage in work (art II-75); protection in the event of unjustified dismissal (art II-90); right of collective bargaining and action (art II-88); fair and just working conditions (art II-91); social partners and autonomous social dialogue (art I-48)¤
	public services¤	Social security and social assistance (art II-94); services of general economic interest (art II-96, art III-122, art III-238)¤
Consequential Argument¤	¤	[meta argument]¤
Democracy Argument¤	¤	[treaty argument]¤
	Democratisation, more powers for EP¤	strenghtening EP (art I-20), e.g. election of President of the Commission (art I-27); codecision ad part of Ordinary legislative procedure (art I-34 + III-396); Union's annual budget (art III-403 - III-409), right to initiative for constitutional amendments (art IV-443 - IV-445) ¤
	subsidiarity¤	National parliaments, early warning mechanisms (art I-11, protocols no 1 + 2); clearer division of competences (art I-12 - I-14)¤
	Citizen Initiative¤	European Citizen Initiative (art I-47)¤
	transparency¤	Founding treaties into one document; transparency principle for EU institutions (art I-50); public Council meeting for deliberation and decision on legislative acts (art I-24 (6))¤
	Double majority¤	Double majority (art I-25)¤

Argument¤
Social Policy Argument¤
Consequential Argument¤
Democracy Argument¤
Institutional Argument¤
Foreign Policy Argument¤
Historical Argument¤
Charter Argument¤
Economic Argument¤
Security Argument¤
Ecological Argument¤

fundamental right to environmental protection (art II-97); ecological horizontal clause (art III-119); polluter-pays-principle (art III-233 (2))¤

4　Comparative EU referendum research

Lists of arguments – yes

Argument	Core statement
YES ARGUMENTS – Vote Yes to the treaty because…	
Charter Argument (ChartA)	…it guarantees the protection of fundamental rights.
Consequentialist Argument (ConsA)	…voting no will have bad consequences for our country.
Democratic Argument (DemA)	…it makes the EU political process more democratic.
Ecological Argument (EcoA)	…it lays the grounds for better environmental protection.
Economic Argument (EA)	…ratification is good for economic development.
Foreign Policy Argument (FPA)	…the EU can play a greater role in the world then.
Historical Argument (HistA)	…our country has benefitted from EU membership.
Institutional Argument (InstA)	…it makes EU institutions work more efficiently.
Security Argument (SecA)	…it promotes judicial and police cooperation in the EU.
Social Policy Argument (SPA)	…it promotes EU social policy.

4 Comparative EU referendum research

Lists of arguments – no

Argument	Core statement
NO ARGUMENTS – Vote no to the treaty because…	
Critique of the Democratic Deficit (DDC)	…it does not remedy the democratic deficit/even increases it.
Economic Critique (EconC)	…it is bad for our economic development.
Ecological Critique (EcoC)	…it is bad for environmental protection.
Enlargement Critique (EnlaC)	…it continues the enlargement agenda instead of stopping it.
Influence Argument (InflA)	…it reduces our country's power in EU decision-making.
Militarization Critique (MilC)	…it leads to a further militarization of the EU.
Critique of Neoliberalism (NeoC)	…it is a manifest of neoliberal ideology.
Plan B Argument (PlanB)	…there is a better option/a Plan B available.
Sanction Vote-Argument (SVA)	…the vote damages our national government.
Socio-ethical Critique (SEC)	…it endangers our national values.
Sovereignty Argument (SovA)	…it brings a further loss of national sovereignty.

4   Comparative EU referendum research

Results - % of documents in which an argument appears at least once
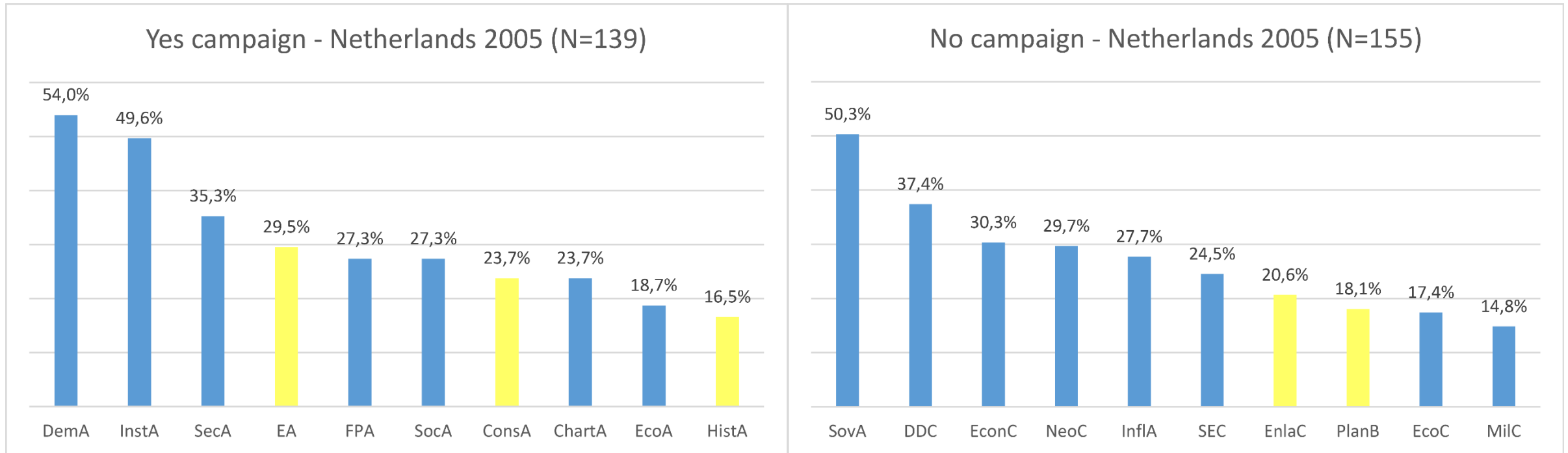
Yes campaign - France 2005 (N=218)

Argument	%
SocA	56,0%
ConsA	54,1%
DemA	49,5%
InstA	45,9%
FPA	44,0%
HistA	36,2%
ChartA	35,3%
EA	25,2%
SecA	14,2%
EcoA	12,8%

No campaign - France 2005 (N=396)

Argument	%
NeoC	74,2%
DDC	40,7%
EconC	32,3%
PlanB	30,1%
SovA	26,5%
MilC	21,5%
EnlaC	14,9%
SEC	14,6%
SVA	0,5%

4   Comparative EU referendum research

Results - % of documents in which an argument appears at least once

Yes campaign - Netherlands 2005 (N=139)

Category	Value
DemA	54,0%
InstA	49,6%
SecA	35,3%
EA	29,5%
FPA	27,3%
SocA	27,3%
ConsA	23,7%
ChartA	23,7%
EcoA	18,7%
HistA	16,5%

No campaign - Netherlands 2005 (N=155)

Category	Value
SovA	50,3%
DDC	37,4%
EconC	30,3%
NeoC	29,7%
InflA	27,7%
SEC	24,5%
EnlaC	20,6%
PlanB	18,1%
EcoC	17,4%
MilC	14,8%

4 Comparative EU referendum research

Results - % of documents in which an argument appears at least once

Yes campaign - Ireland 2008 (N=289)

HistA 41,2% | EA 40,5% | FPA 39,8% | ConsA 39,4% | InstA 34,6% | DemA 29,4% | EcoA 27,0% | ChartA 26,0% | SocA 21,8% | SecA 14,9%

No campaign - Ireland 2008 (N=253)

SovA 53,4% | NeoC 49,4% | DDC 44,3% | MilC 31,6% | PlanB 28,1% | InflA 26,5% | EconC 21,7% | EcoC 8,3% | SEC 7,5% | SVA 1,2%

4   Comparative EU referendum research

Interplay 1: Simple corpus queries: "art." + "article" or "artikel"

	all tokens	all types	art./article	Percent
F 2005	895268	336640	1482	0,17%
F 2005 oui	264763	100436	**299**	**0,11%**
F2005 non	630505	236204	**1183**	**0,19%**
NL 2005	238957	94231	250	0,10%
NL 2005 ja	77392	32004	**23**	**0,03%**
NL 2005 non	161565	62227	**227**	**0,14%**
IRL 2008	465952	157035	571	0,12%
IRL 2008 yes	246376	81752	**155**	**0,06%**
IRL 2008 no	219576	75283	**416**	**0,19%**
IRL 2009	401215	153494	663	0,17%
IRL 2009 yes	193498	74889	**191**	**0,10%**
IRL 2009 no	207717	78605	**472**	**0,23%**

4 Comparative EU referendum research

Preliminary conclusions

- Some support for each of the hypotheses 1 and 2, as cases present a mixed picture of narrow issue-orientation measured by treaty references campaigners made and orientation towards broader aspects of EU integration.

- In most of the cases, treaty-arguments seem to be the most frequent type of arguments.

- No support for H3 (second-order-orientation) on the level of argumentation/campaigning. Idea of using the referendum as a sanction-vote against the incumbent government is mostly not made explicit. If it is, then more so in France and Ireland than in the Netherlands.

- No support for H4. Findings stand in clearer contrast to H4, the narrower issue-orientation is defined.

- Overall picture: Pro-treaty campaigners tend to rely more on meta-arguments than treaty opponents that tend to more narrowly refer to the treaty at hand ("chercher des phrases").

- This does not come with any assessment of the substantial quality of arguments (sound or misleading).

- However, it seems to be more difficult for campaigners in EU treaty referendums to stick to the document at hand, which might be another indicator for structural advantages of treaty opponents in EU referendums.

Topic Modelling with R topicmodels (WasN$_{BTW17}$-Corpus, AfD)

Topic Nr.	Top-Terms	Thema
1	afd sagen wähler links deutschland grün kommen schauen nein rot bleiben bringen gleich sehen welt sitzen mensch bundestag heissen gehen	X
2	deutsch lernen sprache kultur sprechen verstehen deutsche schule arabisch gut geschichte integration warum viele heute vielleicht damals integrieren wegen türkisch	Integration über Sprache
3	deutschland problem jahr thema deutsch mensch sicherheit viele prozent sozial neu million bringen merkel asylante bekommen lösen sehen schaffen gehen	Merkels Flüchtlingspolitik
4	europa land usa europäsch deutschland frankreich krieg russland afrika million welt soros polen merkel jahr regierung staat plan stehen grenze	Länder und Personen
5	nazi afd links partei demokratie recht meinung rechts politisch rechtrechts demokratisch bezeichnen antifa medium leute beschimpfen mitglied nennen meinungsfreiheit heute	Unfaire Anfeindungen
6	afd partei höcke petry herr gauland gut wÄhler frauke stehen meuthen alternative mitglied rede gehen wahl sagen björn npd wahlkampf	AfD-Politiker und was sie so machen
7	kommen gut voll gehen schön halten paar bekommen raus jahr richtig wissen gar neu viele rein bestimmt falsch vielleicht lang	X
8	brauchen darlehen haus erhalten dringend zeit groß gut com bauen kaufen gmail schnell klein antwort mögen geschäft name sicher rufen	Baufinanzierung
9	deutschland staat merkel deutsch grenze mensch recht natülich kommen alternative asyl sehen land politik angela klar europäisch international gesetz volk	Merkels Flüchtlingspolitik
10	deutschland euro staat politik wirtschaft geld welt groß gut schuld jahr neu nehmen dafür griechenland schaden stark letzt sozial bereits	Kritik an „Schuldenunion"

5 Experimental application of (Structural) Topic Modelling

Lack of categorical clarity

Gilbert Ryle: The Concept of Mind, University of Chicago [1949] 2002.

of a division.
One more illustration. A foreigner watching his first game of cricket learns what are the functions of the bowlers, the batsmen, the fielders, the umpires and the scorers. He then says 'But there is no one left on the field to contribute the famous element of team-spirit. I see who does the bowling, the batting and the wicket-keeping; but I do not see whose role it is to exercise *esprit de corps.*' Once more, it would have to be explained that he was looking for the wrong type of thing. Team-spirit is not another cricketing-operation supplementary to all of the other special tasks. It is, roughly, the keenness with which each of the special tasks is performed, and performing a task keenly is not performing two tasks. Certainly exhibiting team-spirit is not the same thing as bowling or catching, but nor is it a third thing such that we can say that the bowler first bowls and then exhibits team-spirit or that a fielder is at a given moment either catching or displaying esprit de corps.

This book first arose out of a passage in Borges, out of the laughter that shattered, as I read the passage, all the familiar landmarks of my thought – our thought, the thought that bears the stamp of our age and our geography – breaking up all the ordered surfaces and all the planes with which we are accustomed to tame the wild profusion of existing things, and continuing long afterwards to disturb and threaten with collapse our age-old distinction between the Same and the Other. This passage quotes a 'certain Chinese encyclopaedia' in which it is written that 'animals are divided into: (a) belonging to the Emperor, (b) embalmed, (c) tame, (d) sucking pigs, (e) sirens, (f) fabulous, (g) stray dogs, (h) included in the present classification, (i) frenzied, (j) innumerable, (k) drawn with a very fine camelhair brush, (l) *et cetera*, (m) having just broken the water pitcher, (n) that from a long way off look like flies'. In the wonderment of this taxonomy, the thing we apprehend in one great leap, the thing that, by means of the fable, is demonstrated as the exotic charm of another system of thought, is the limitation of our own, the stark impossibility of thinking *that*.

Michel Foucault: The Order of Things, Psychology Press 2002, Preface, p. xvi.

5 Experimental application of (Structural) Topic Modelling
Fundamental problems with topic modelling

- ## Context problem
  What is the optimal scope of co-occurrence?

- ## Problem of replicability (and evaluation)
  How can we make TM results replicable?

- ## Lack of categorical clarity/interpretive analytics
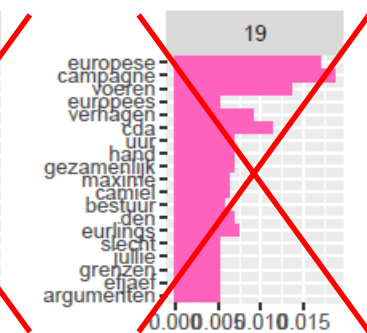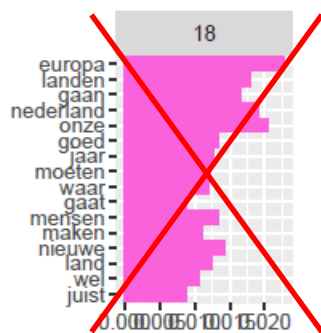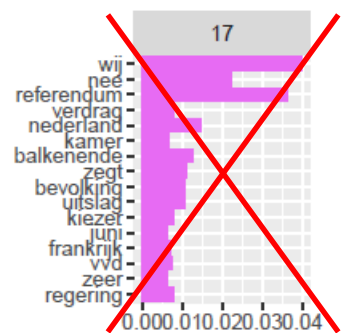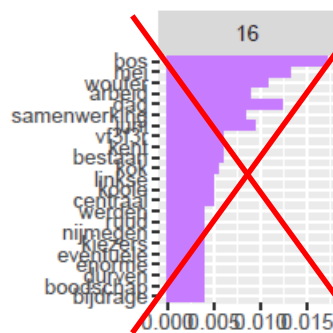 What is a topic in Topic Modelling? An unavoidable risk of producing category errors?

  (cf. Brookes & McEnery 2019)
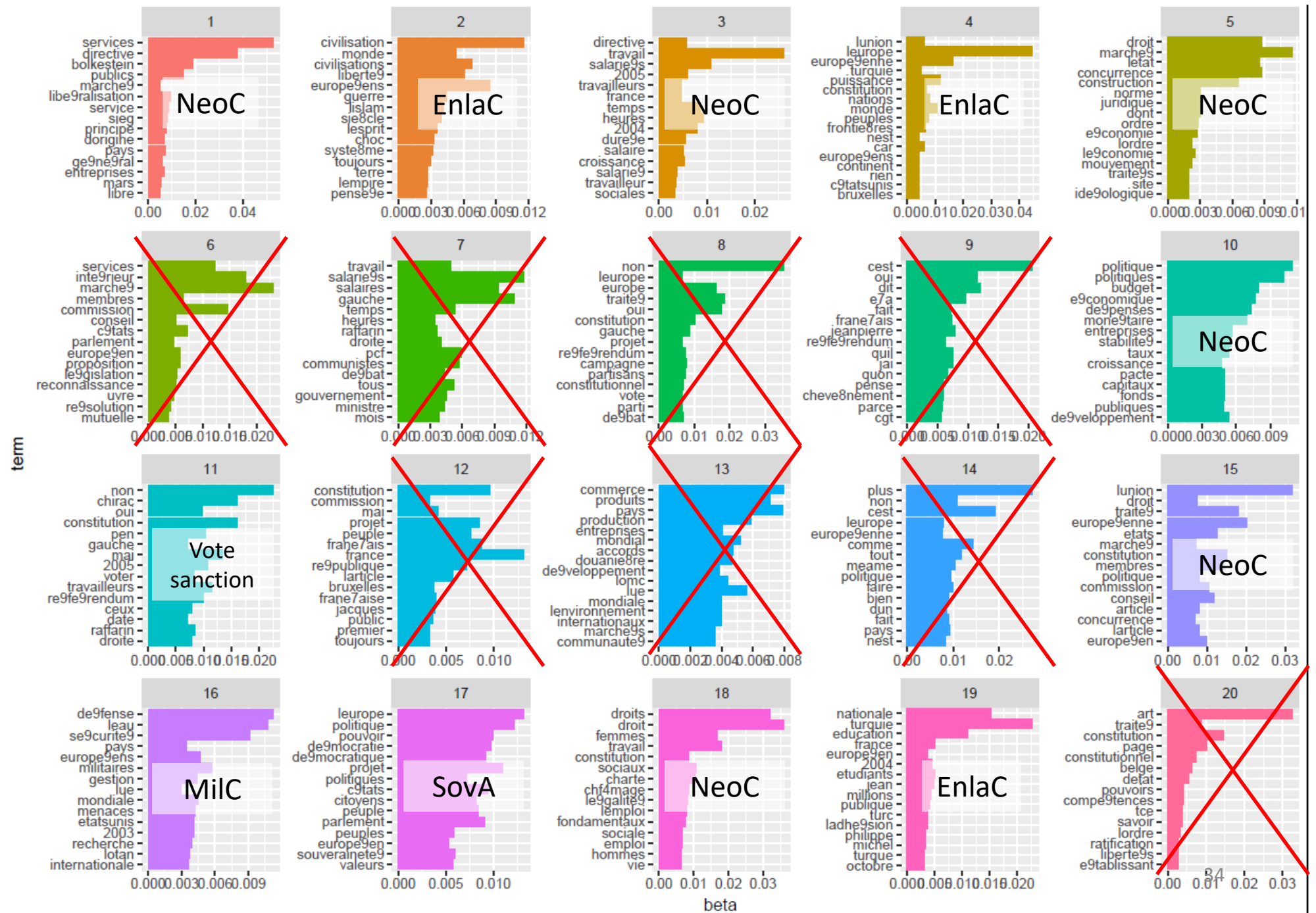
Brookes, G., & McEnery, T. (2019). The utility of topic modelling for discourse studies: A critical evaluation. *Discourse Studies*, *21*(1), 3–21.
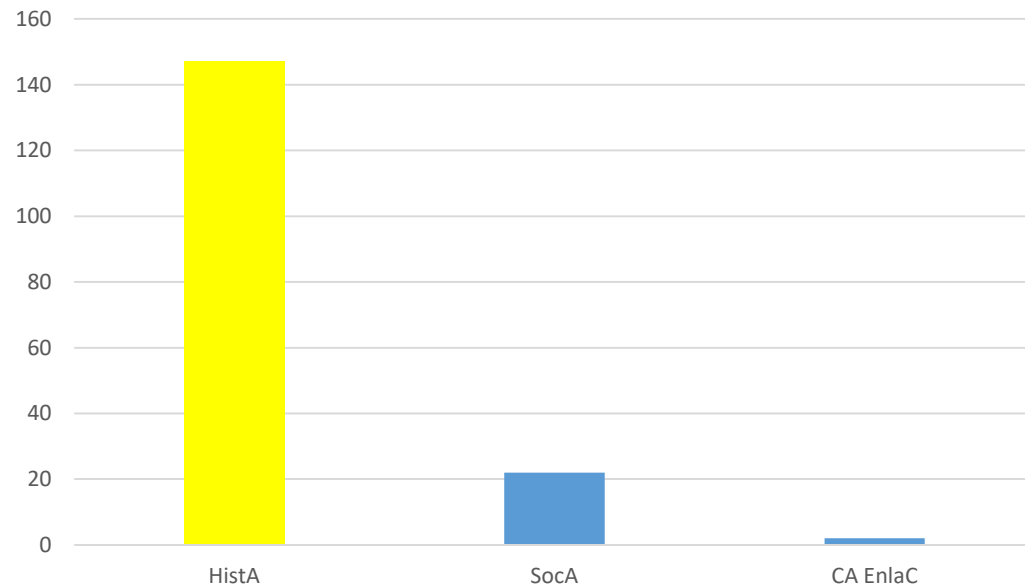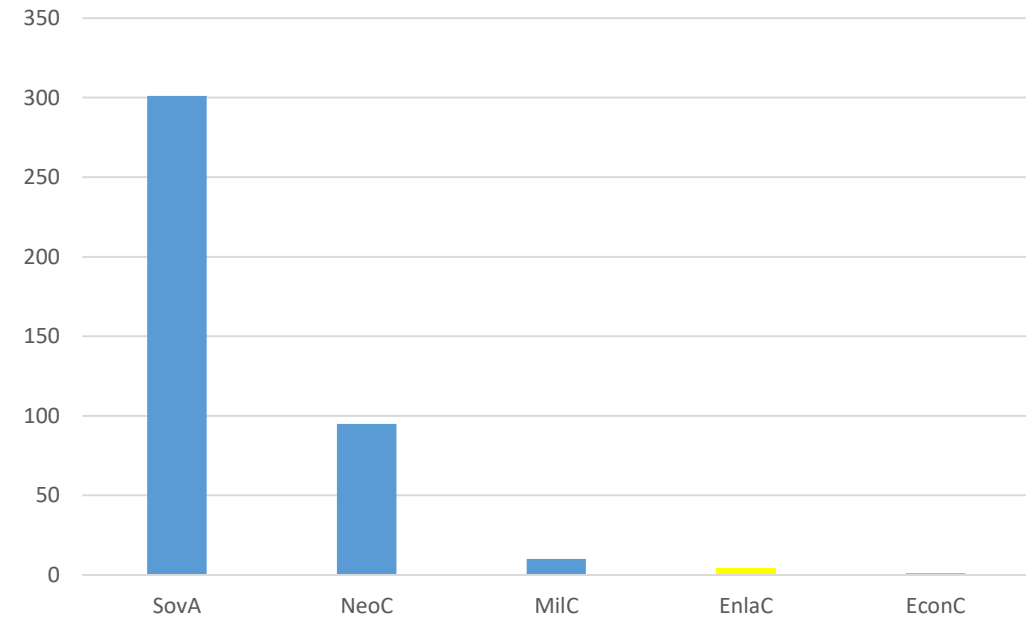
Ja Netherlands
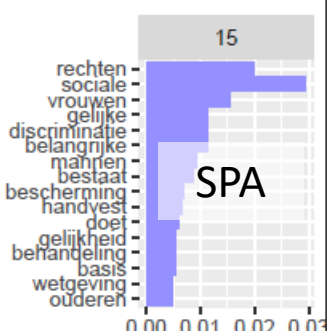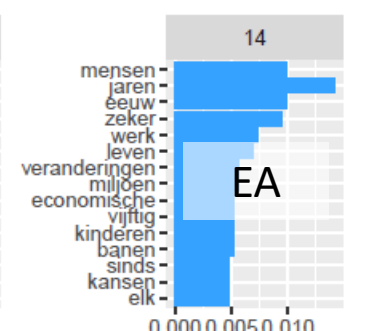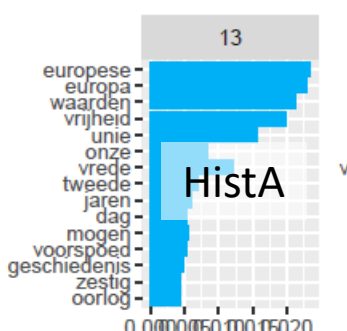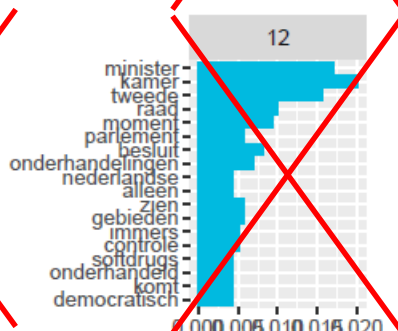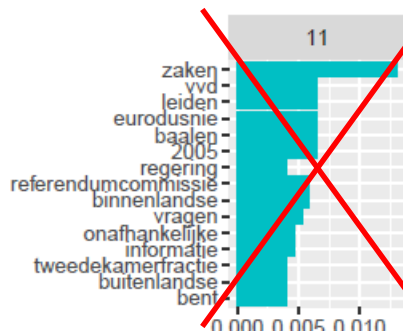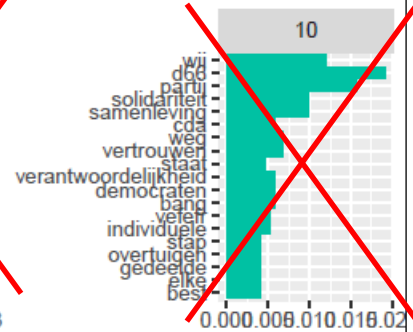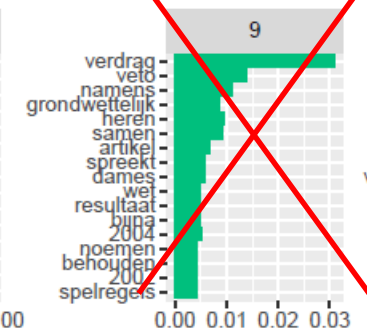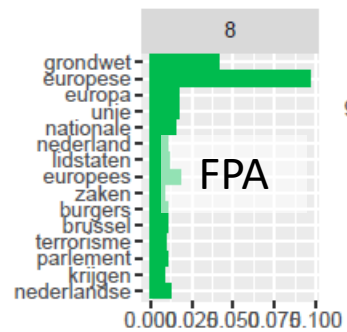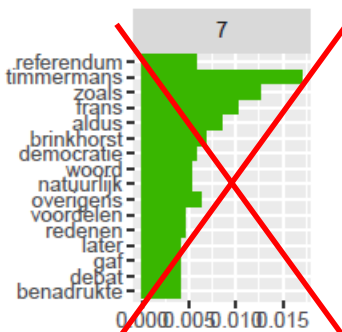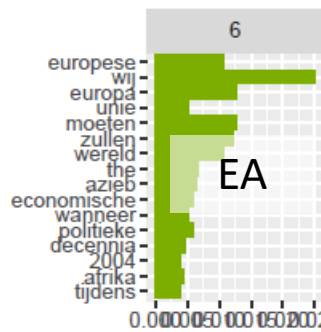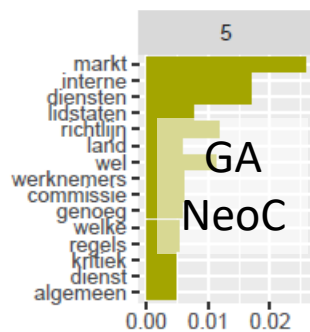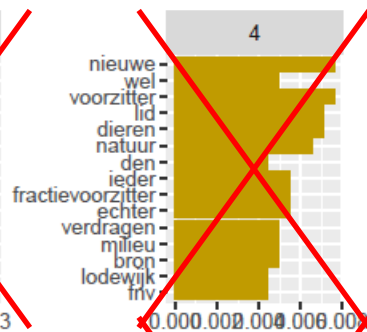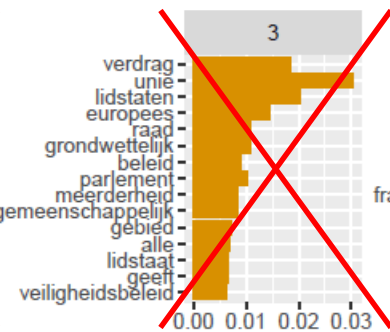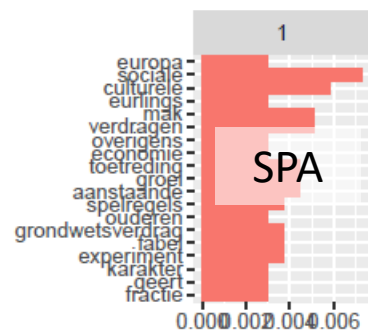
Experimental application of (Structural) Topic Modelling

Topic Modelling with R topicmodels (EU-Ref Corpora) – France

Topic probabilities (gamma) - France 2005 - oui

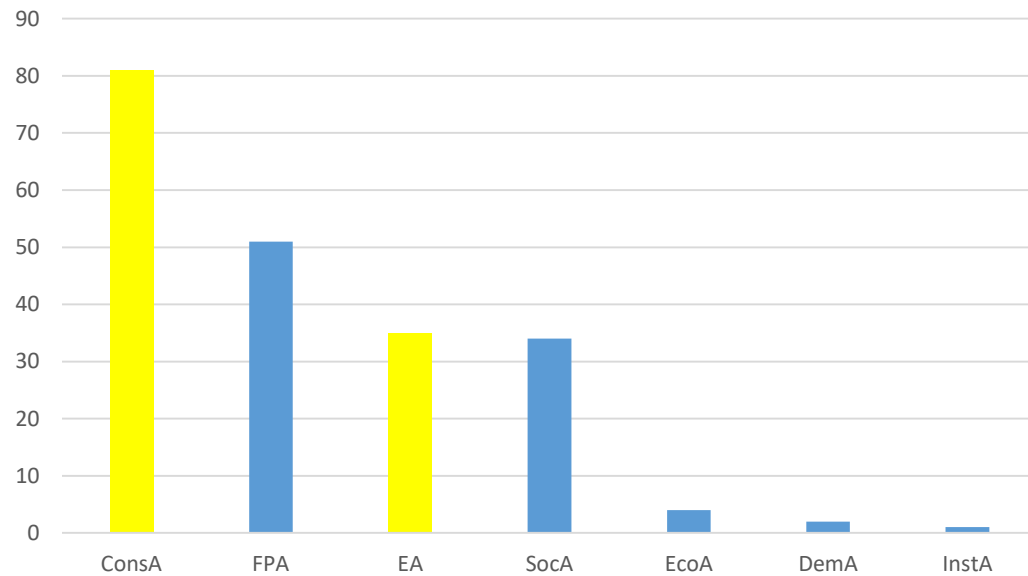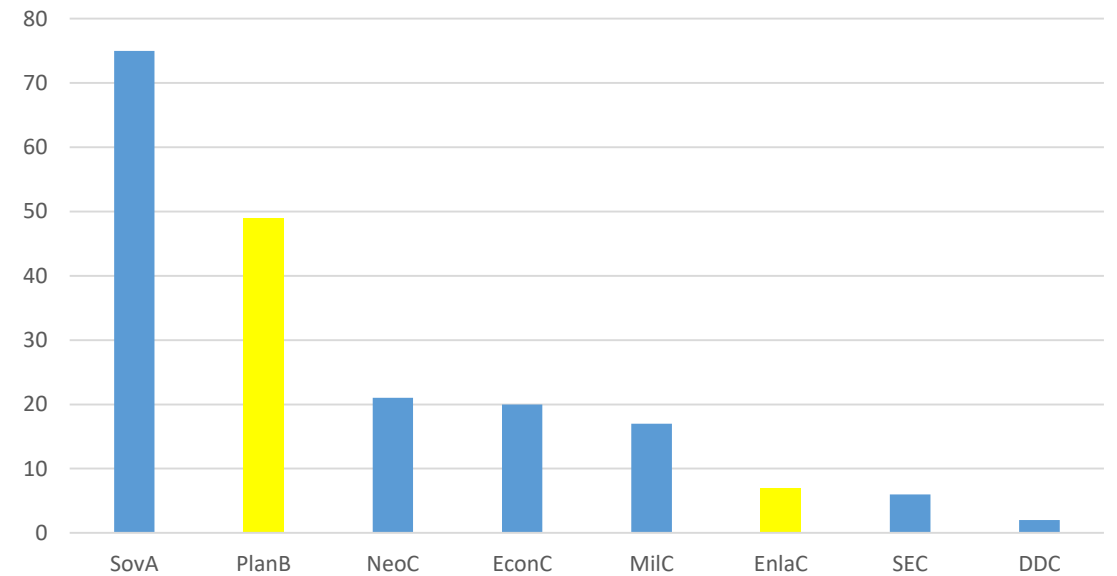Topic probabilities (gamma) - France 2005 - non

Comparative EU referendum research

Topic Modelling with R topicmodels (EU-Ref Corpora) – Ireland 2008

Topic probabilities (gamma) - Ireland 2008 yes

Topic probabilities (gamma) - Ireland 2009 no

5   Comparative EU referendum research

Structural Topic Modelling

- Developed by Margaret E. Roberts et al., R package STM (Roberts et al. 2014, 2016, 2019)
- Some crucial advantages regarding semantic coherence and replicability such as setting options for the initial seed for the LDA algorithm, built-in iteration and evaluation functions to be called.
- MOST IMPORTANT: Option to include co-variates, enrichment of model with meta-data, possibility to measure effects

Roberts, M. E., Stewart, B. M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S. K., Rand, D. G. (2014). Structural Topic Models for Open-Ended Survey Responses. American Journal of Political Science, 58(4), 1064–1082. Roberts, M. E., Stewart, B. M., & Airoldi, E. M. (2016). A Model of Text for Experimentation in the Social Sciences. Journal of the American Statistical Association, 111(515), 988–1003. Roberts, M. E., Stewart, B. M., & Tingley, D. (2019). stm : An R Package for Structural Topic Models. Journal of Statistical Software, 91(2).

5 Comparative EU referendum research

Structural Topic Modelling

```
Topic 1 Top Words:
 EcoA     Highest Prob: ireland, treaty, irish, will, business, lisbon, europe, vote, yes, economic
          FREX: ibec, chamber, commerce, business, exports, association, jobs, chambers, exporters, investment
          Lift: auctioneers, cif, counterproductive, donations, exported, exposure, manufacturing, muchneeded, performs
, retail
          Score: chamber, commerce, exports, exporters, ibec, butler, association, ihf, chambers, fdi
Topic 2 Top Words:
 CAMP     Highest Prob: treaty, lisbon, people, will, irish, europe, profit, campaign, vote, public
          FREX: profit, boyd, barrett, socialist, privatisation, militarisation, postal, wing, posted, antiwar
          Lift: boyd, marie, terror, ambulances, bare, coconvenor, conned, convenor, dgra, disturb
          Score: boyd, profit, barrett, privatisation, gregory, smyth, richard, ailbhe, antiwar, postal
Topic 3 Top Words:
 CAMP     Highest Prob: campaign, treaty, vote, party, lisbon, said, cfir, will, yes, people
          FREX: posters, greene, cfir, poster, print, advertising, canvass, green, source, leaflets
          Lift: bernard, erected, fatal, luas, tube, bebo, clips, daddy, damages, deutsche
          Score: cfir, greene, posters, richard, volunteers, poster, print, spokesman, canvass, poll
Topic 4 Top Words:
 MilC     Highest Prob: military, irish, defence, treaty, will, states, european, member, nato, constitution
          FREX: eda, nato, superstate, structured, operations, battlegroups, shannon, renamed, military, capabilities
          Lift: allegiance, ben, berlinplus, commonality, contravention, danes, expenditures, irishman, irishmen, junck
er
          Score: nato, eda, battlegroups, renamed, military, twentysix, operations, pana, iraq, defence
Topic 5 Top Words:
 SovA     Highest Prob: treaty, european, states, new, union, lisbon, member, will, national, council
          FREX: teu, parliaments, citizenship, prime, federal, amended, laws, postlisbon, tfeu, comment
          Lift: abolishing, advent, alsodefine, andensure, arerepresentatives, attaching, ave, beintergovernmental, cam
eras, changeart
          Score: teu, tfeu, art, postlisbon, citizenship, presidents, federal, amended, parliaments, prime
Topic 6 Top Words:
 CAMP     Highest Prob: treaty, lisbon, irish, vote, referendum, will, people, government, ireland, europe
          FREX: mcdonald, morgan, lou, begg, siptu, ictu, congress, votenoie, mary, class
          Lift: butit, chairpersons, feedback, gratitude, leinen, abstentions, advising, betters, caolain, frightening
          Score: mcdonald, siptu, morgan, begg, lou, ictu, votenoie, rossa, caolin, congress
Topic 7 Top Words:
```

Comparative EU referendum research

Structural Topic Modelling – Estimate Effects

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.06865    0.01635   4.200 3.16e-05 *** SPA
campyes      0.02691    0.02156   1.248    0.212
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Topic 16:

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.01033    0.01205   0.857    0.392
campyes      0.09308    0.01781   5.226 2.55e-07 *** CAMP
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Topic 17:

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.10114    0.01379   7.332 9.09e-13 *** EconC/SEC
campyes     -0.08256    0.01753  -4.711 3.20e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Topic 18:

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.06434    0.01147   5.610 3.34e-08 *** EconC
campyes     -0.04955    0.01527  -3.246  0.00125 **
```
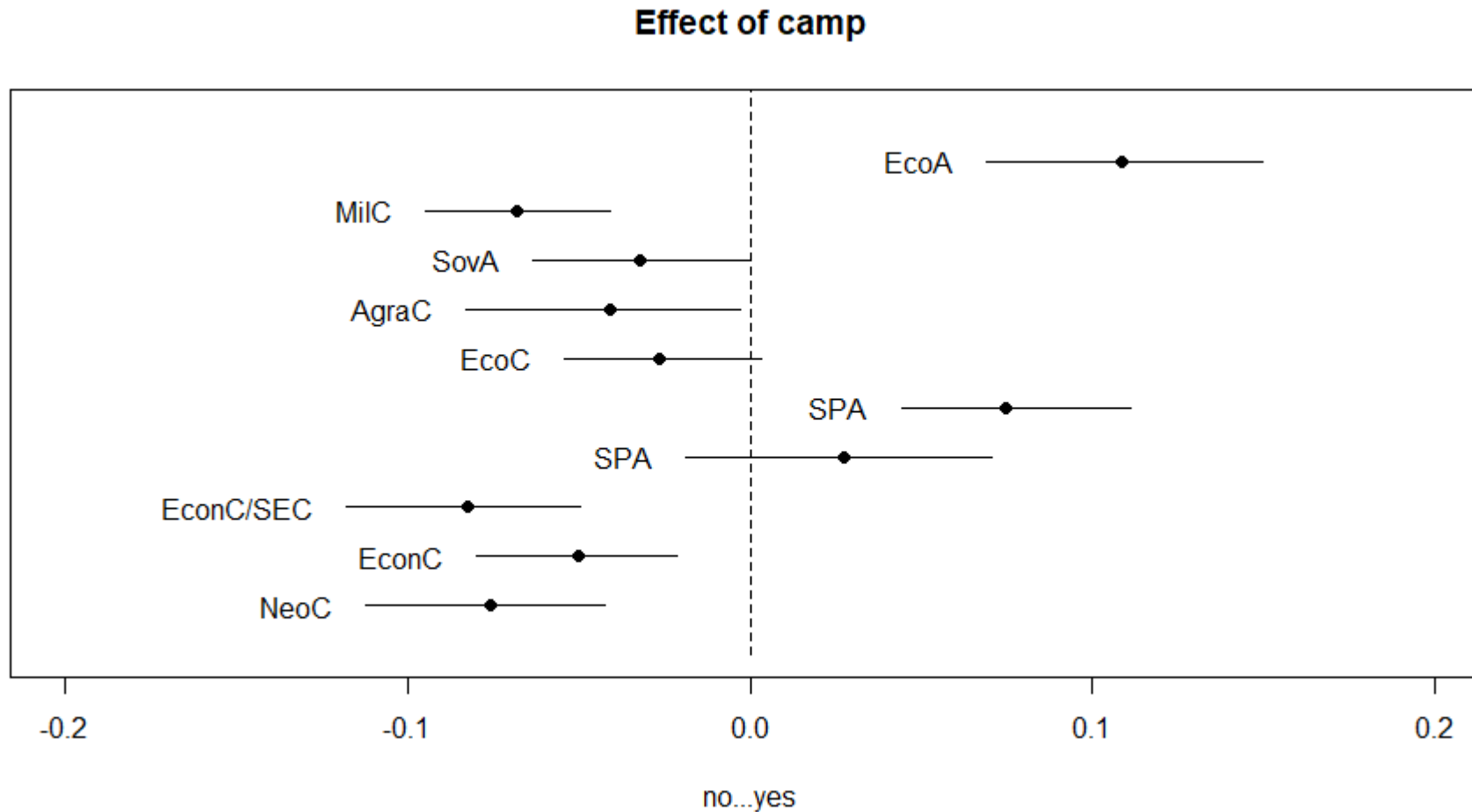
Comparative EU referendum research

Structural Topic Modelling – Estimate Effects

Effect of camp

Conclusions

- Computational methods are helpful to answer meaningful political science questions and thus contribute to actual research interests and debates of the discipline.

- In example 1, a network analysis of EP candidates' interaction lends support to the second-order hypothesis but also helps to reveal and explain some variation in transnational linkages.

- In example 2, qualitative research provides most substantial findings.

- Simple LDA application with an "eyeballing" approach to the necessary interpretative work is not helpful.

- However, a combination with corpus linguistic measures and machine learning approaches like topic modelling might be helpful to augment research to larger datasets for the same research object.

- Moreover, discourse analysis with STM helps to identify discursive patterns and allows for the estimation of effects of pre-defined co-variates. This functionality pretty much serves the expectations of mainstream political science with hypotheses derived from theory.