
Computational Social Science

Course #04199, module 04IN2042

How can we analyze political processes
on the web?

Markus Strohmaier

Prof. Dr.

WeST / IWVI U. of Koblenz-Landau, Germany

e-mail: markus.strohmaier@gesis.org

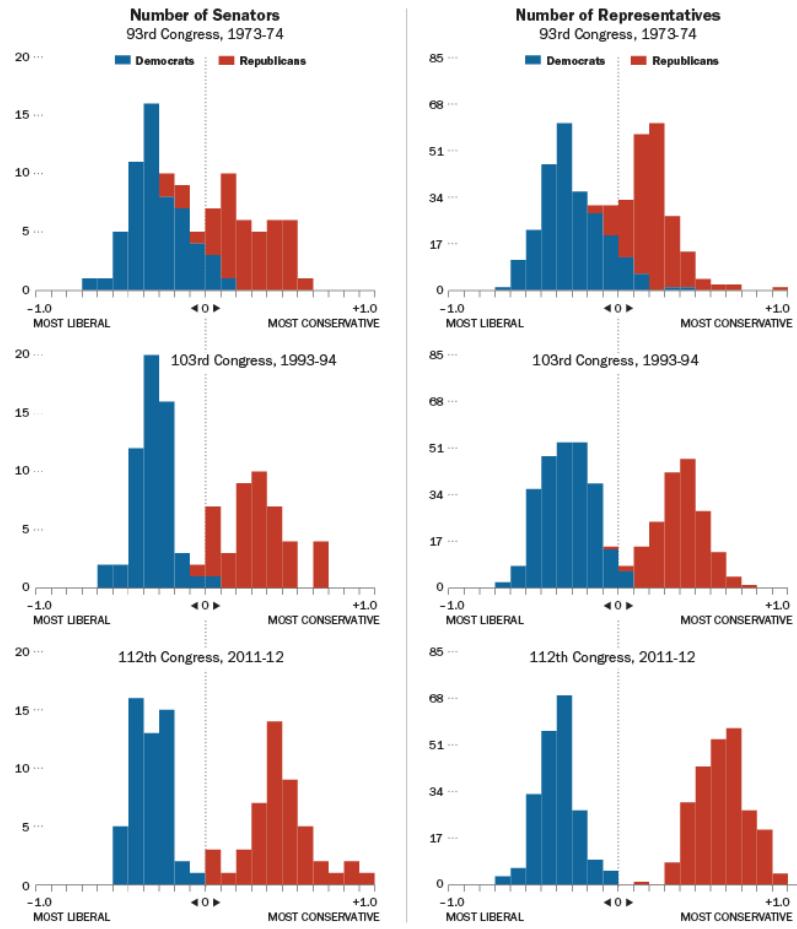
web: <http://www.markusstrohmaier.info>

In part based on slides by C. Kling and C. Wagner

Polarization in US Congress 1973-2012

In Congress as Well as Public, the Center Increasingly Cannot Hold

Ideological scores of senators and representatives based on roll-call votes. Negative numbers represent liberal views and positive numbers conservative views



<http://www.pewresearch.org/fact-tank/2014/06/12/polarized-politics-in-congress-began-in-the-1970s-and-has-been-getting-worse-ever-since/>

Sources: Royce Carroll, Jeff Lewis, James Lo, Nolan McCarty, Keith Poole and Howard Rosenthal, Voteview.com

PEW RESEARCH CENTER

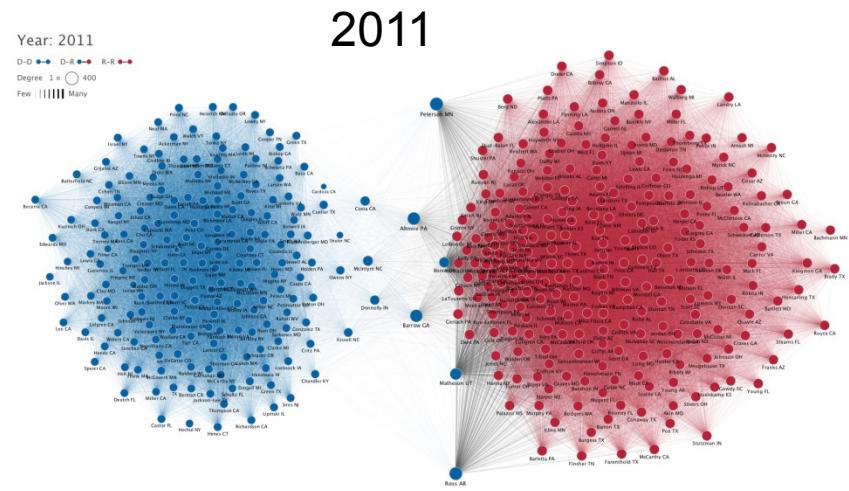
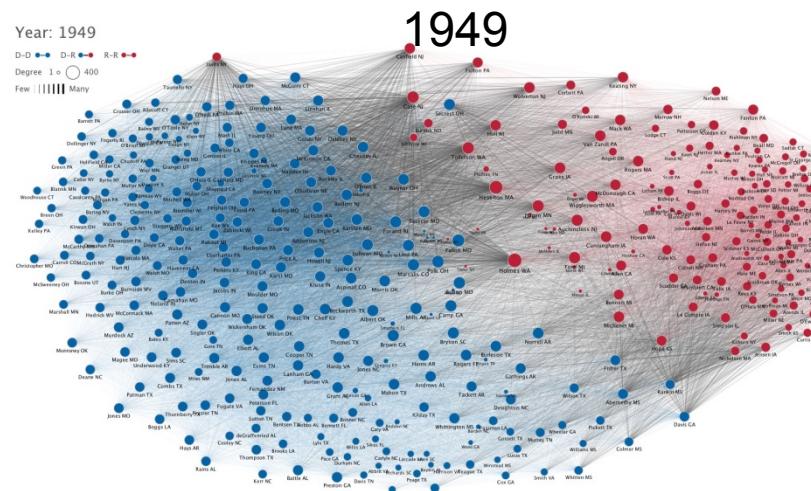
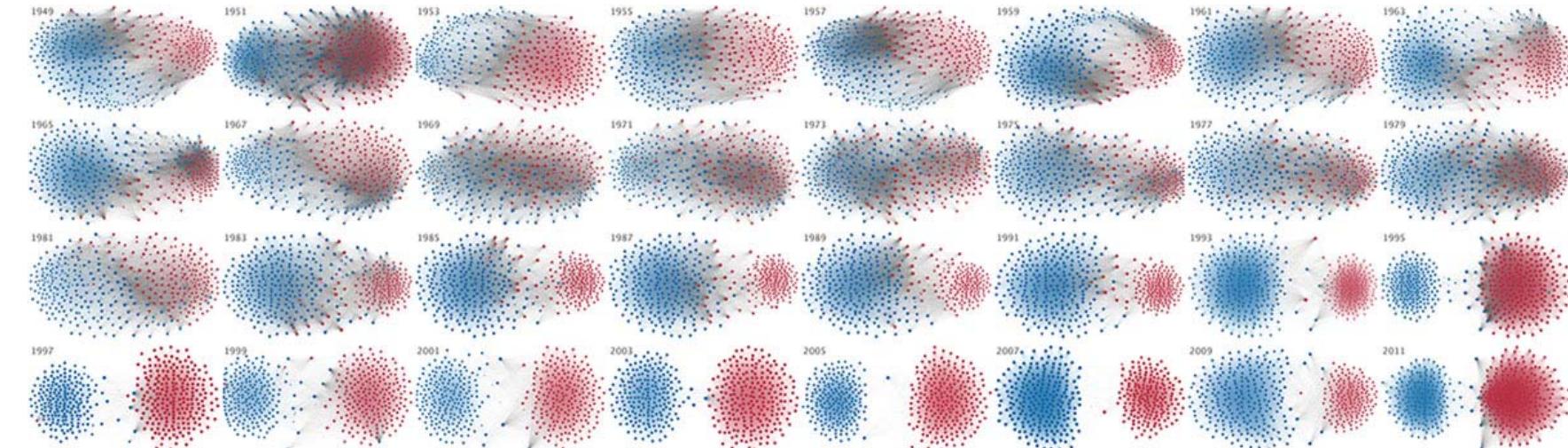
article by [Clio Andris](#), David Lee, Christian E. Gunning, John A. Selden, Mauro Martino and Marcus J. Hamilton.
data visualization by [Mauro Martino \(January, 2014\)](#)

related sources:

Andris, C. et al (2013) santa fe institute working paper (nov. 11, 2013)

Andris, C. (2011) doctoral dissertation, mit, chapter 5

Polarization in US Congress 1949-2011

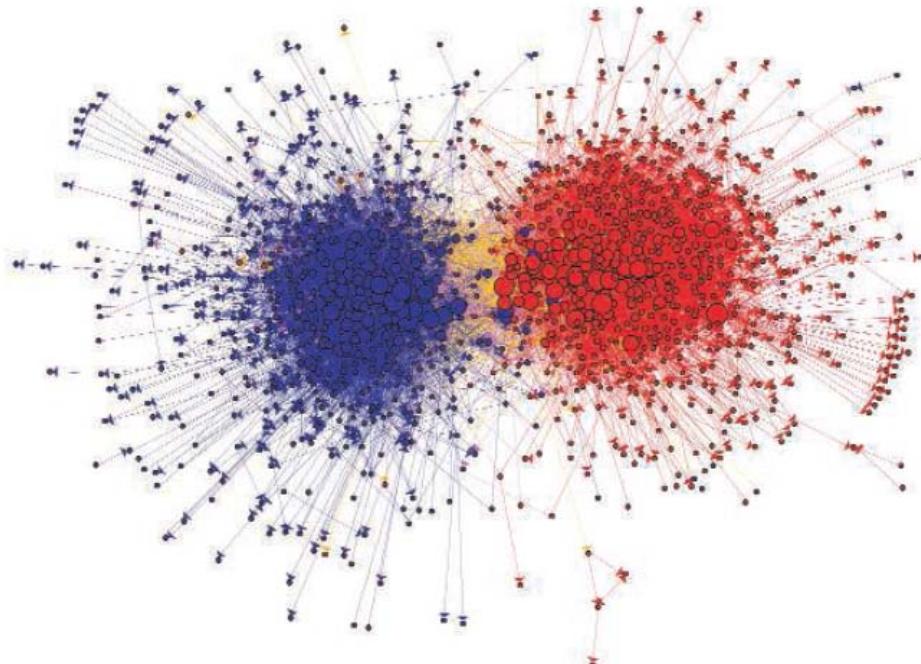


Markus Strohmaier

2015

http://www.mamartino.com/projects/rise_of_partisanship/

Polarization in US Political Weblogs 2004



Data from the blogosphere. Shown is a link structure within a community of political blogs (from 2004), where red nodes indicate conservative blogs, and blue liberal. Orange links go from liberal to conservative, and purple ones from conservative to liberal. The size of each blog reflects the number of other blogs that link to it. [Reproduced from (8) with permission from the Association for Computing Machinery]

Lazer, David, et al. "Life in the network: the coming age of computational social science." *Science (New York, NY)* 323.5915 (2009): 721.

Political Misinformation

Bogus Grass-Roots Politics on Twitter

Data-mining techniques reveal fake Twitter accounts that give the impression of a vast political movement.

By Kurt Kleiner on November 2, 2010



How true? This network graph shows the connections between 6,278 accounts that used the hashtag #gop in September and October 2010.

Researchers have found evidence that political campaigns and special-interest groups are using scores of fake Twitter accounts to create the impression of broad grass-roots political expression. A team at Indiana University used data-mining and network-analysis techniques to detect the activity.

"We think this technique must be common," says [Filippo Menczer](#), an associate professor at Indiana University and one of the principal investigators on the project. "Wherever there are lots of eyes looking at screens, spammers will be there; so why not with politics?"

The research effort is dubbed the [Truthy project](#), a reference to comedian Stephen Colbert's coinage of the word "truthiness," or a belief held to be true regardless of facts or logic. The goal was to uncover organized propaganda or

<http://Truthy.indiana.edu/>

Today's agenda

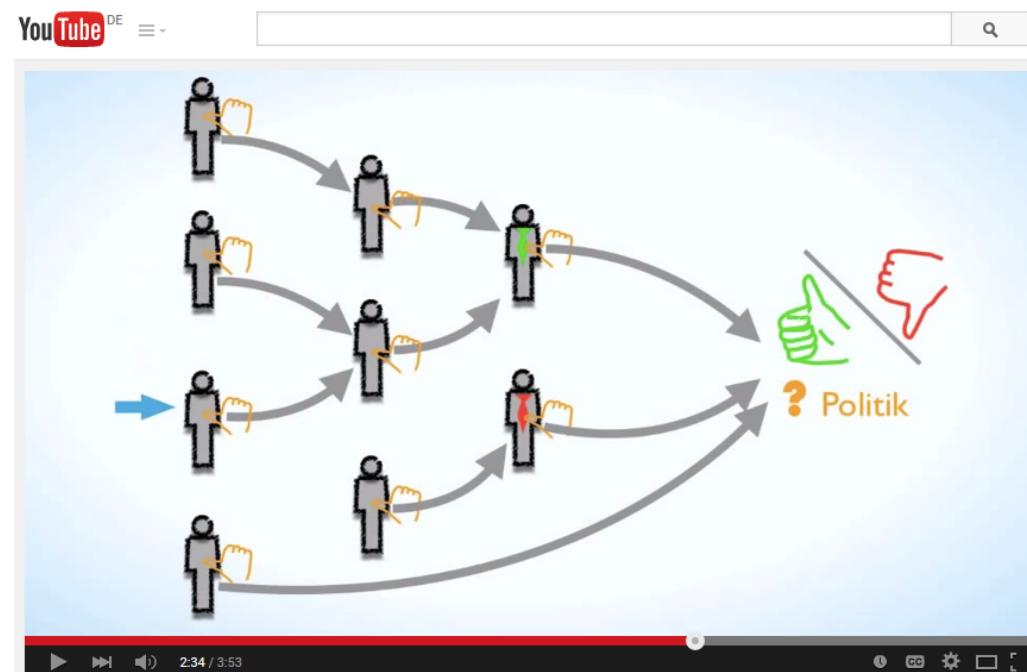
2 examples of Computational Political Science:

1. The Pirate Party and LiquidFeedback as a platform for political processes
2. The German National Elections 2013 on Twitter

Based on slides/research done in collaboration with C. Kling, J. Kunegis, H. Hartmann, S. Staab, H. Lietz, C. Wagner and A. Bleier

Liquid Feedback and Germany's Pirate Party

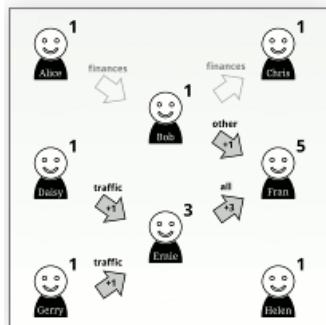
Liquid Democracy



https://www.youtube.com/watch?v=r0G_vuWTOUw

LiquidFeedback

LiquidFeedback is more than Liquid Democracy



Liquid Democracy

Scalability through division of labor
The basic idea is a democratic system in which most issues are decided (or strongly suggested to representatives) by direct referendum. Considering nobody has enough time and knowledge for every issue, votes can be dynamically delegated by topic. Delegations are transitive and can be revoked at any time. Liquid Democracy is sometimes referred to as »Delegated« or »Proxy Voting«.



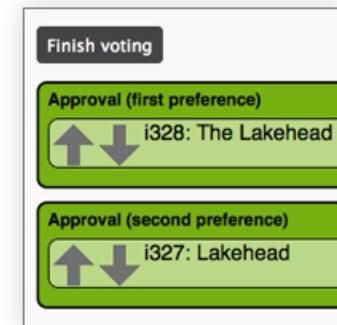
Collective Moderation

Proportional representation of minorities
LiquidFeedback doesn't rely on a request commission and doesn't need a moderator. Instead, all participants gain equal rights in a scalable structured discussion process where it is ensured that minorities gain a fair share of representation and that even individuals may put up their proposals for discussion. The system is designed in such way that noisy minorities won't harm other minorities in the discussion process.



Fully Transparent Decision Process

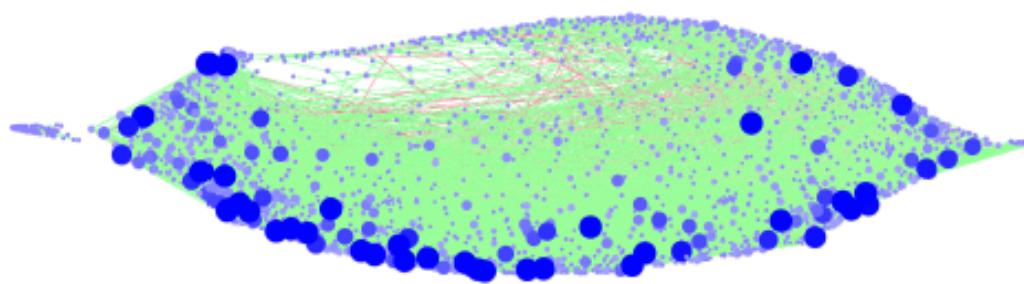
Protection against non-transparent lobbying
Predefined rules and timings ensure that plans on decision processes are made public in time. Decisions are made by recorded vote only, and all voting-relevant data in LiquidFeedback is made available to all participants in both human- and machine-readable form. This enables a transparent decision-making process and ensures that participants can verify the voting procedure.



Preferential Voting

Equal treatment of competing alternatives
LiquidFeedback doesn't ask predefined questions but encourages participants to suggest alternatives. A sophisticated voting system is facilitated to allow participants to express their opinions without necessity of tactical considerations. Its mathematical properties avoid vote-splitting and allow that similar proposals don't harm each other.

The German Pirate Party and LiquidFeedback



(a) Delegation network

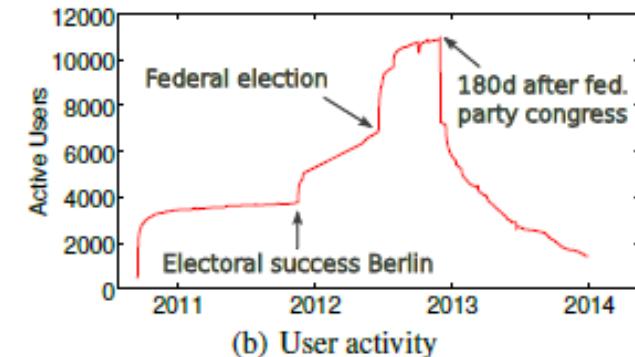


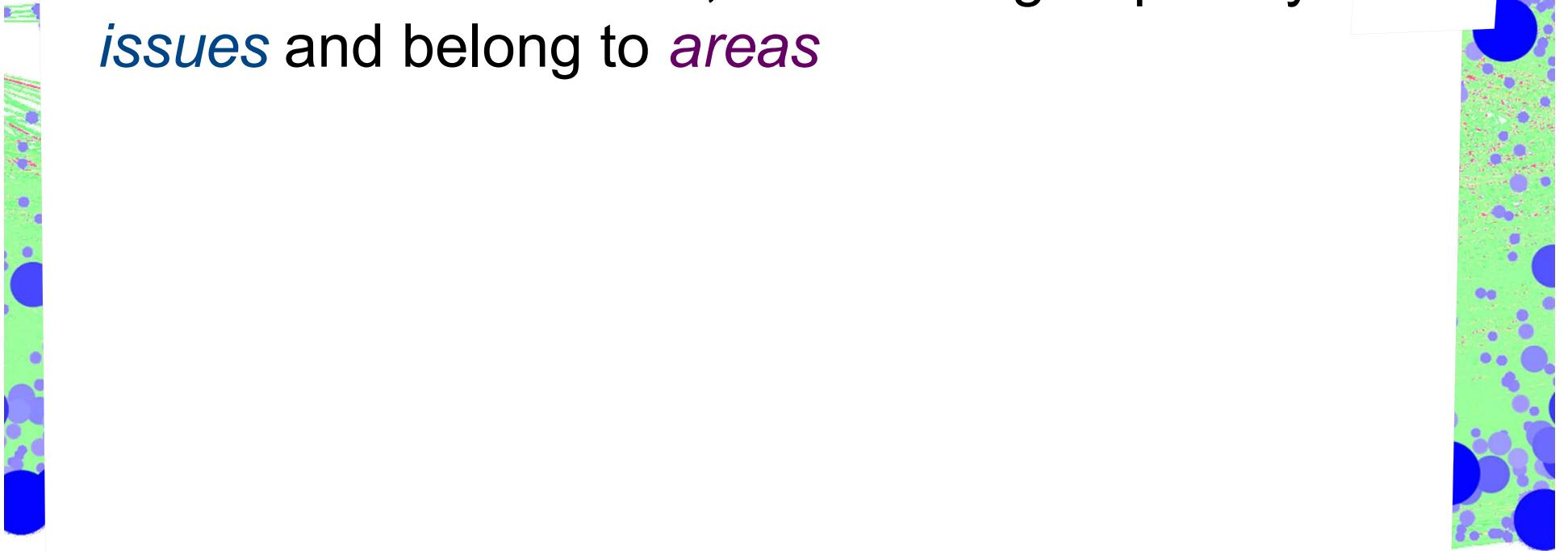
Figure 1: (a) The delegation network of the LiquidFeedback dataset. Obtained from the delegative democracy platform of the German Pirate Party. Red edges denote removed users. The layout was computed in (i) the emergence of a large connected component of delegations. (b) User activity over time. Users are labelled inactive if they have not voted.

Observation period	2010/08/13 – 2013/11/25
Votes	499,009
Users	13,836
Delegations	14,964
Proposals (<i>Initiatives</i>)	6,517
Issues	3,565
Areas	22

Table 1: The LiquidFeedback dataset. Obtained from the delegative democracy platform of the German Pirate Party (e Party member), edges denote delegations. Red edges denote removed users. The layout was computed in (i) the emergence of a large connected component of delegations. (b) User activity over time. Users are labelled inactive if they have not voted. The number of delegations received by other users is denoted by the matrix. We can observe *super-voters*, i.e. voters that have received many delegations. The user activity of the German Pirate Party over time shows a sharp increase followed by a decrease in activity.

LiquidFeedback – German Pirate Party

Users create *initiatives*, which are grouped by *issues* and belong to *areas*



LiquidFeedback – German Pirate Party

Users create *initiatives*, which are grouped by *issues* and belong to *areas*

Area:

Environmental issues

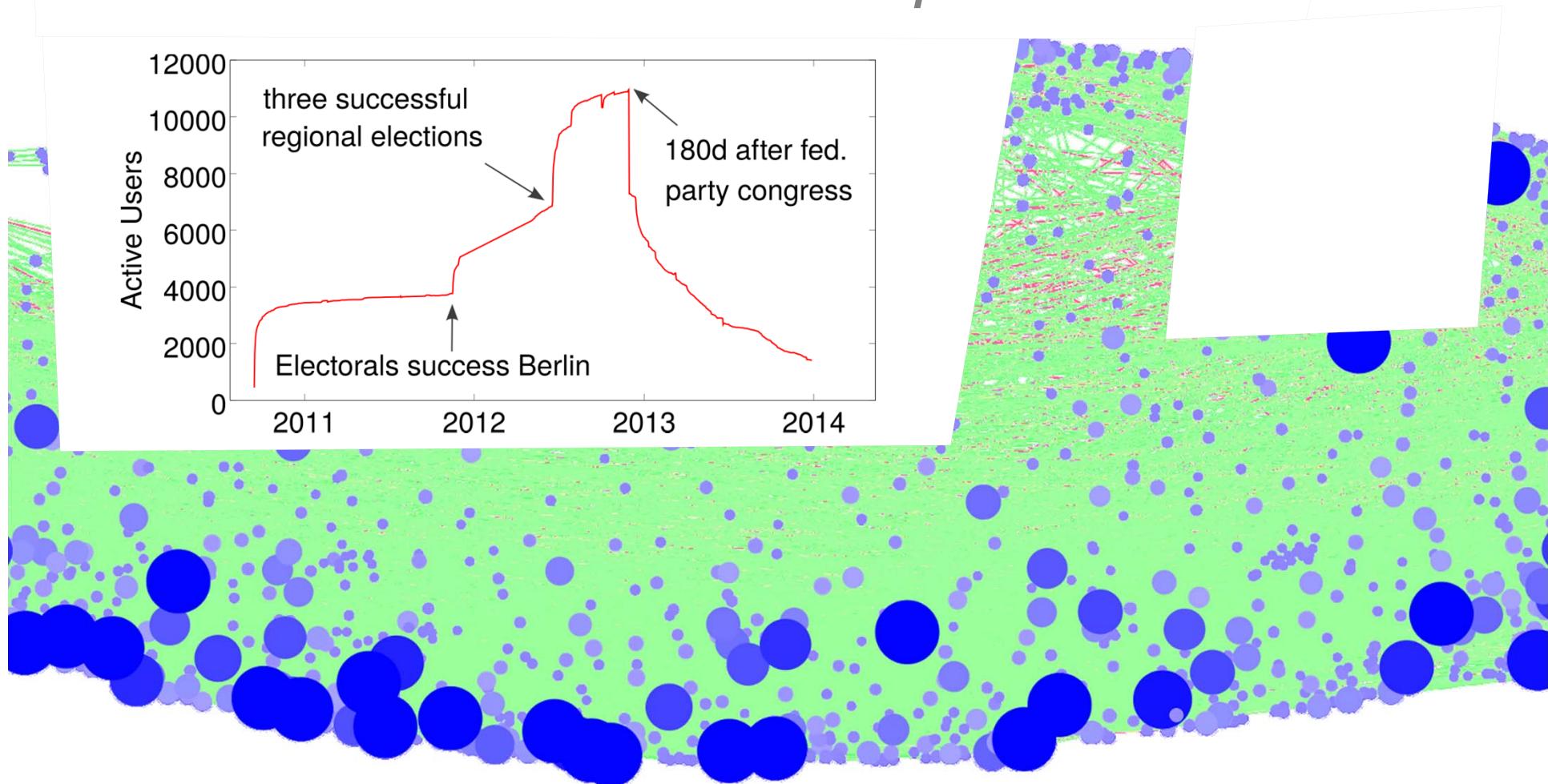
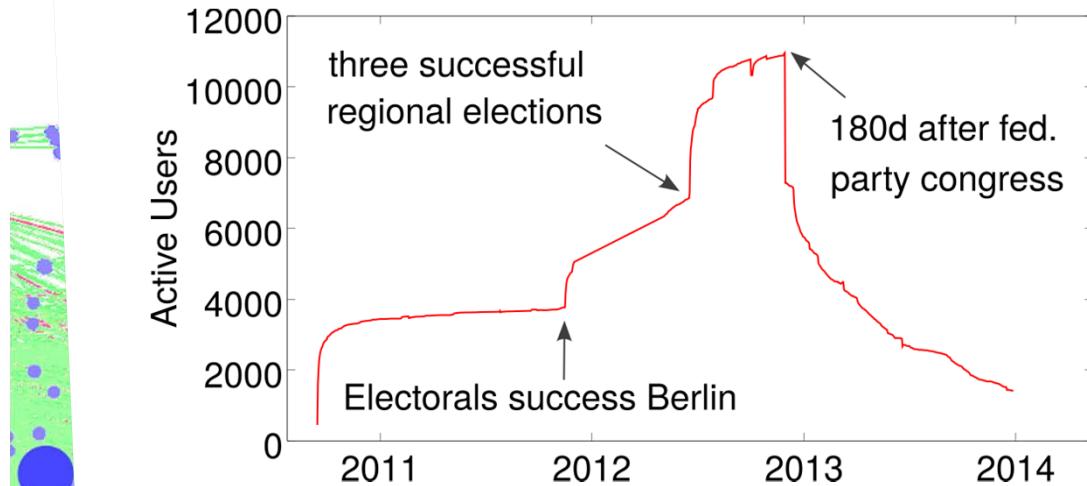
Issue:

CO2 output has to be reduced.

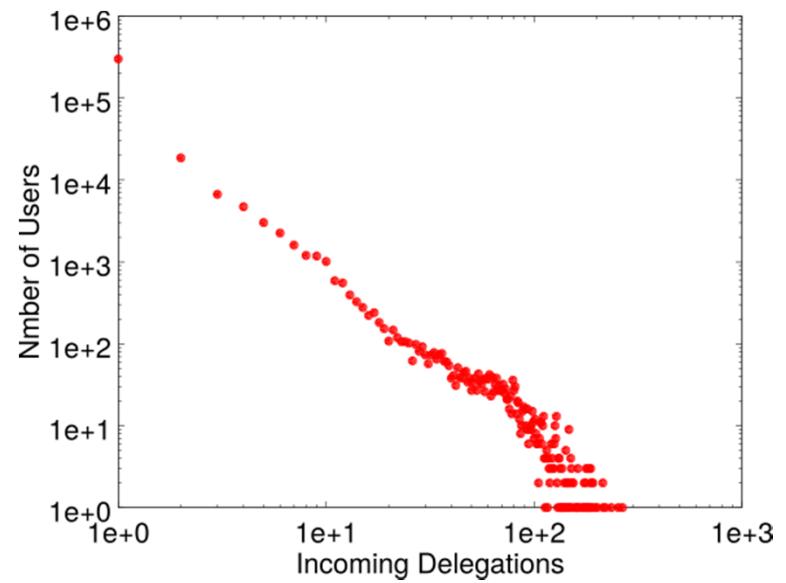
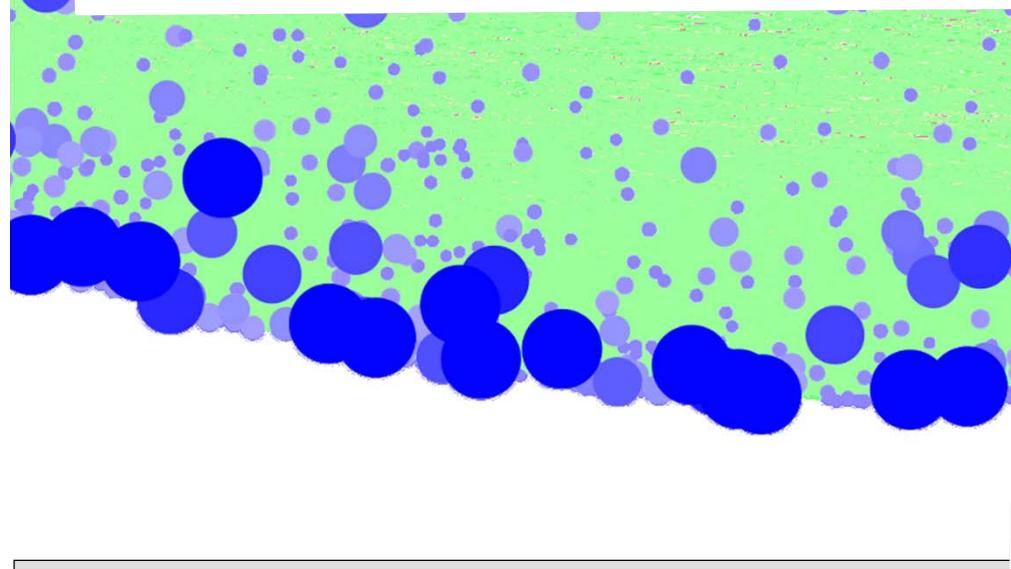
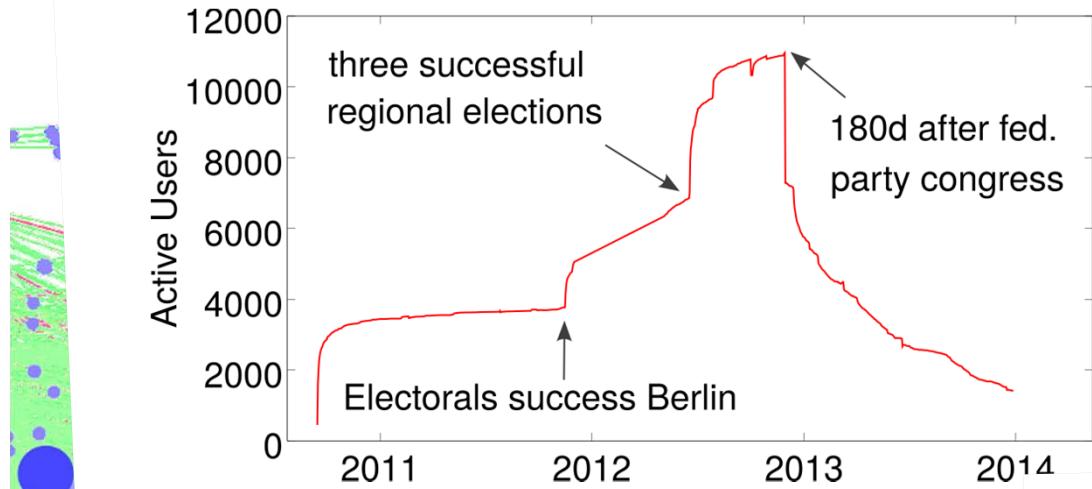
Initiative:

Subsidise wind turbines!

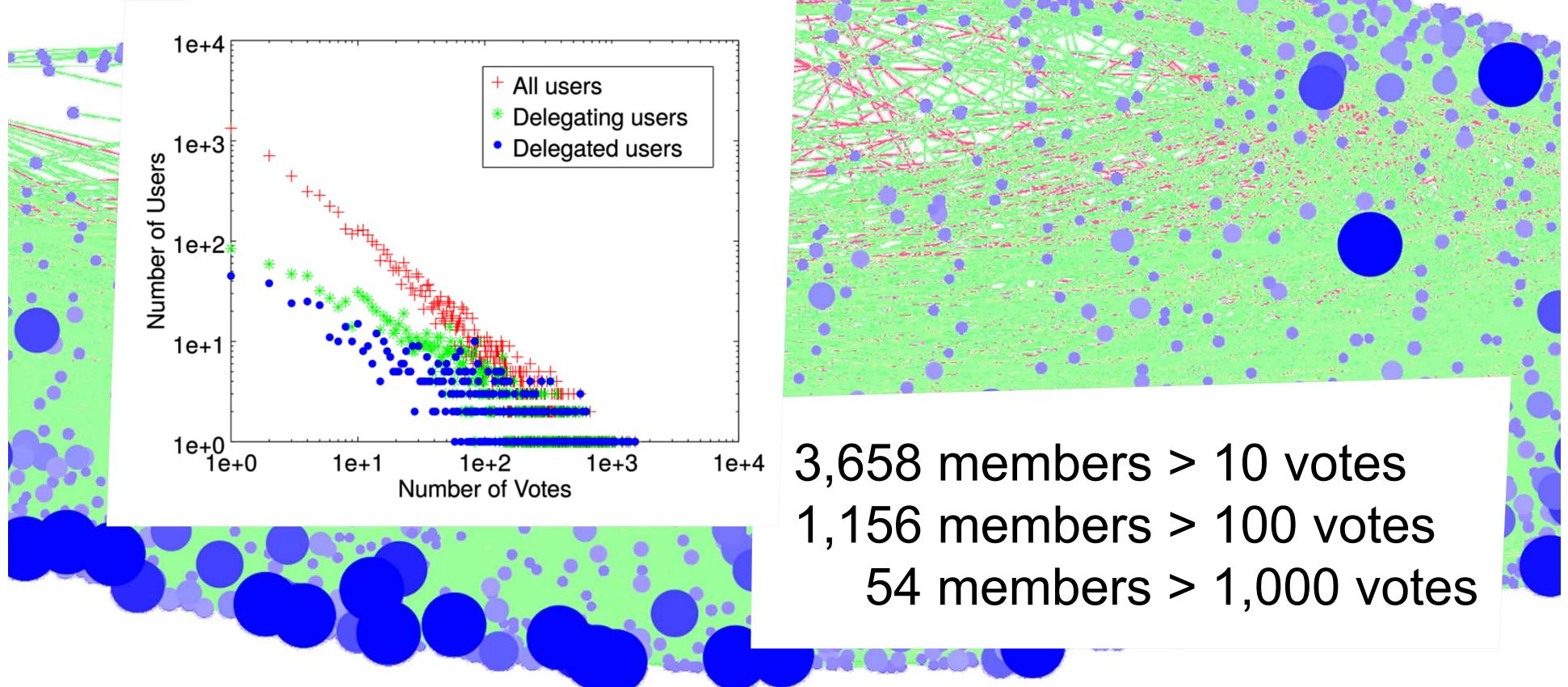
Dataset – First Impressions



Dataset – First Impressions



Dataset – First Impressions



LiquidFeedback – German Pirate Party

Users create *initiatives*, which are grouped by *issues* and belong to *areas*

Area:

Environmental issues

Issue:

CO2 output has to be reduced.

Initiative:

Subsidise wind turbines!

Delegations on *global*, *initiative*, *issue* and *area level*

→ “Back-delegations” possible

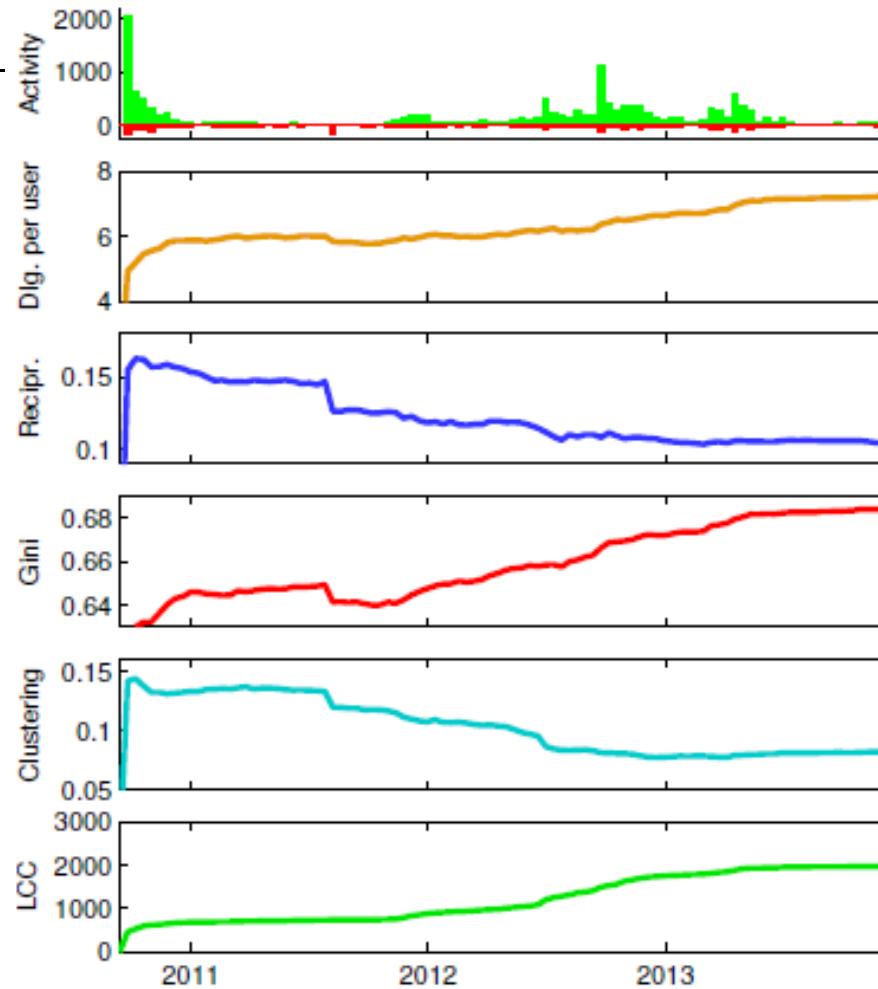
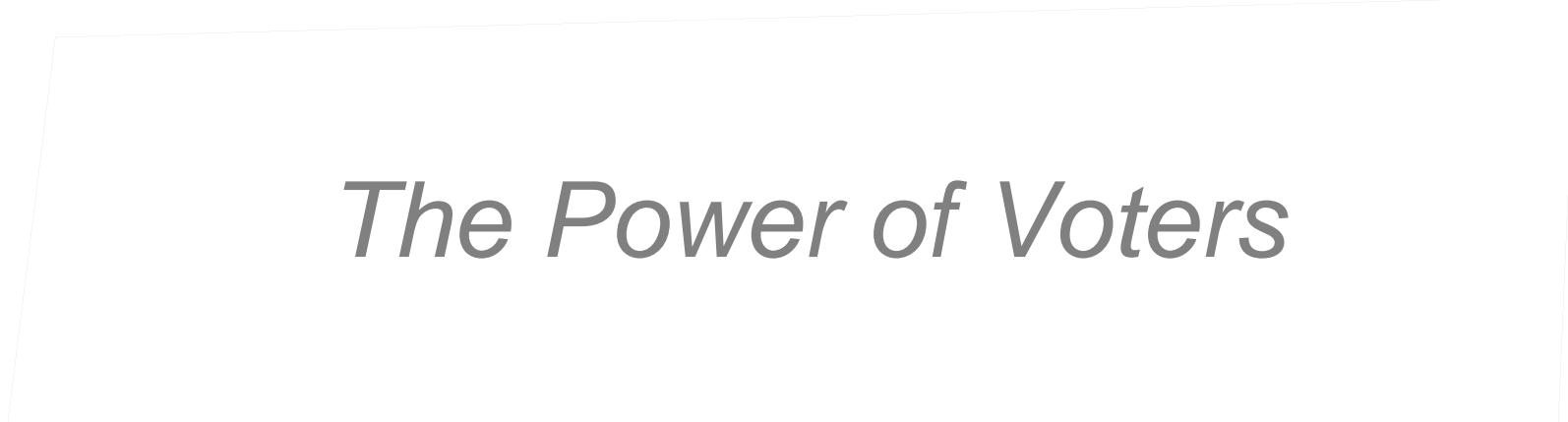


Figure 5: Changes in the delegation network. From top to bottom, we show added and removed delegations, changes in the per-user delegation count, inequality of incoming delegations measured by the Gini coefficient and the reciprocity which gives the proportion of mutual delegations. Note that mutual delegations are only permitted for distinct areas or *issues*.



The Power of Voters

Power

Ability to influence the outcome of a vote

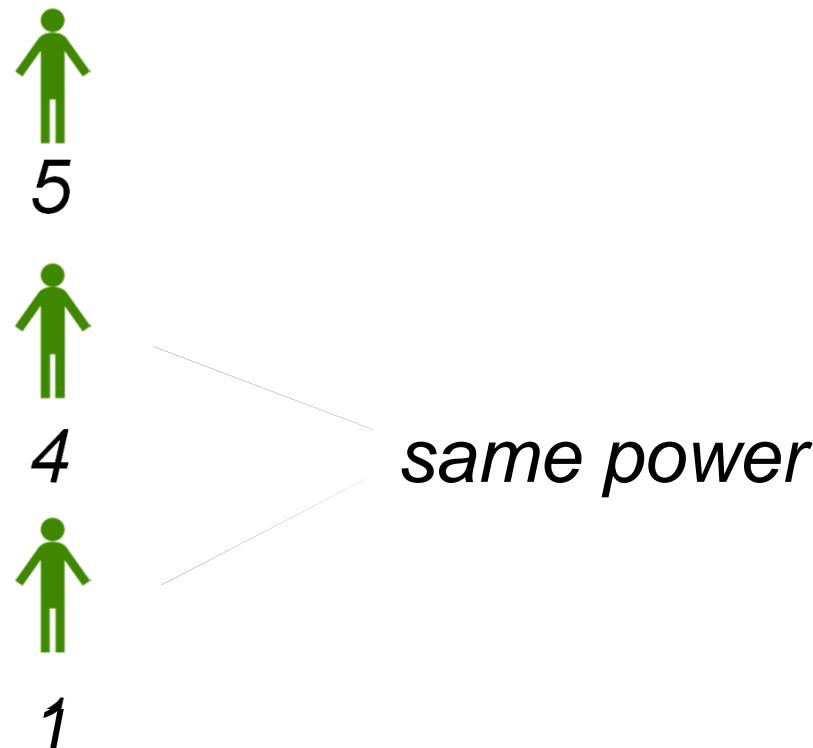
Power

Ability to influence the outcome of a vote



Power

Ability to influence the outcome of a vote



Power Indices

*Given voting weights of all voters in a vote:
Predict the probability that a given user will be
able determine the outcome of a vote*

*Banzhaf power index:
Votes are independent*

*Shapley power index:
Votes are homogeneous*

Power

Banzhaf power index:

Votes are independent

Shapley power index:

Votes are homogeneous

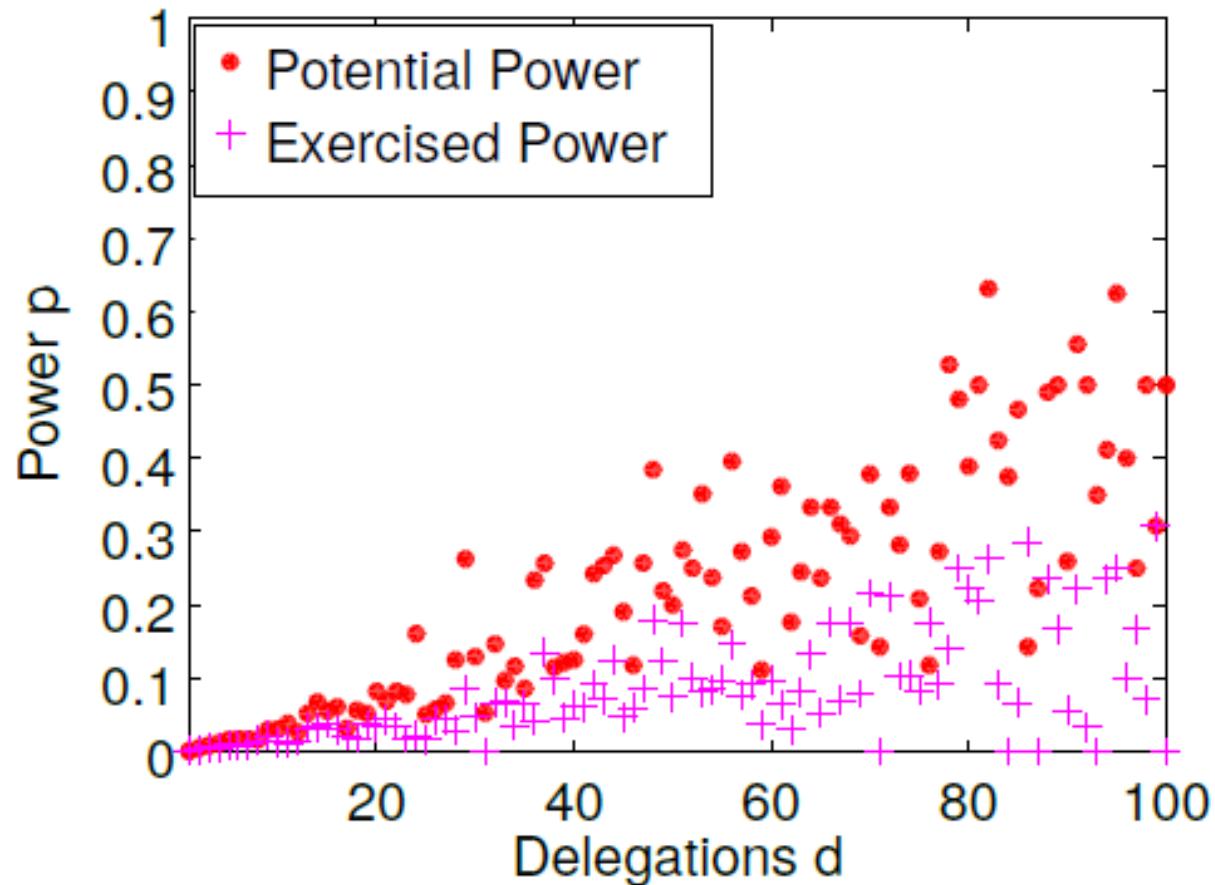
Potential Power:

Measured power in the dataset

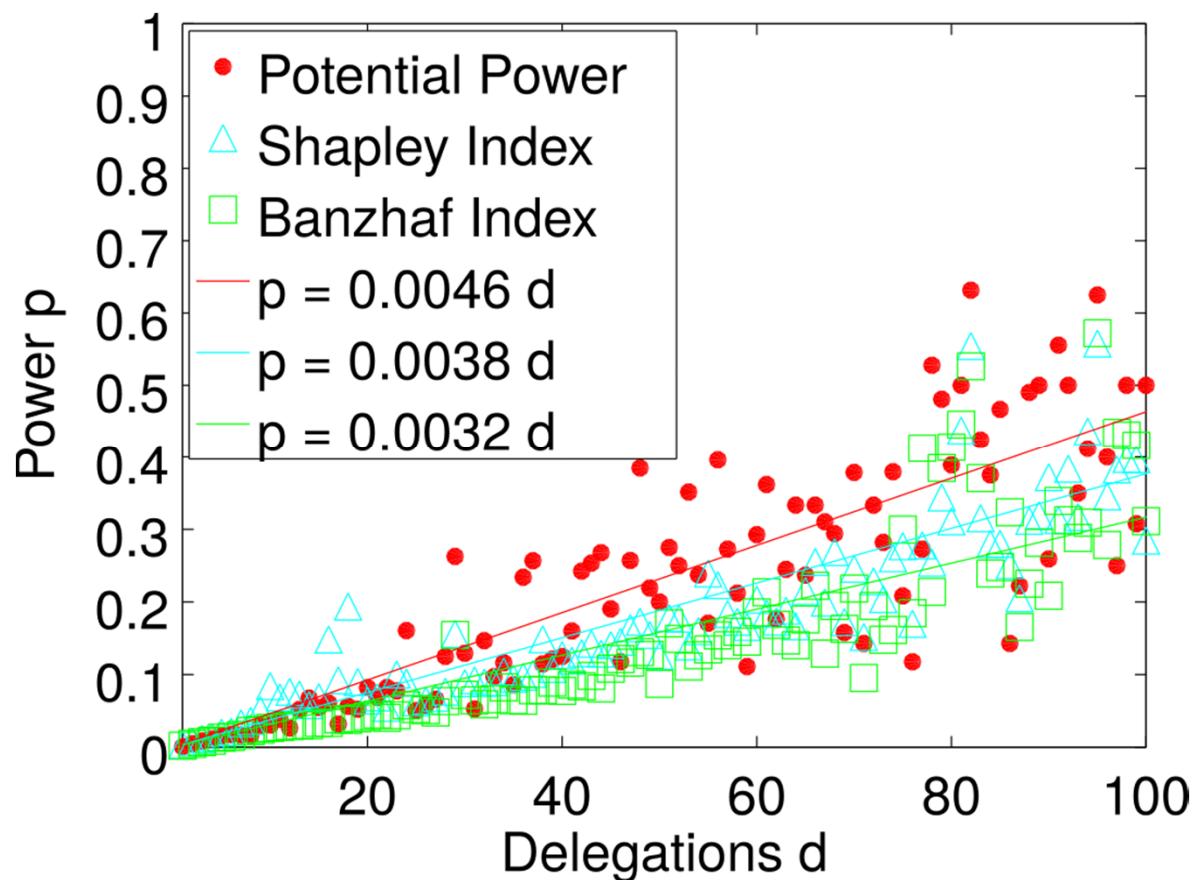
Exercised Power:

Power used to actually turn votes

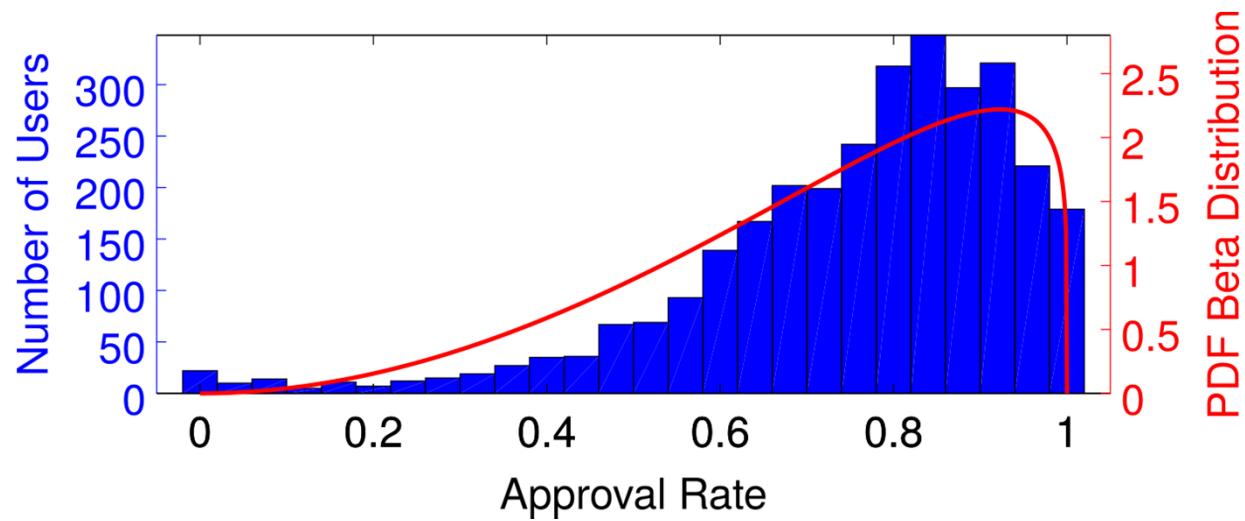
Power Indices



Power Indices

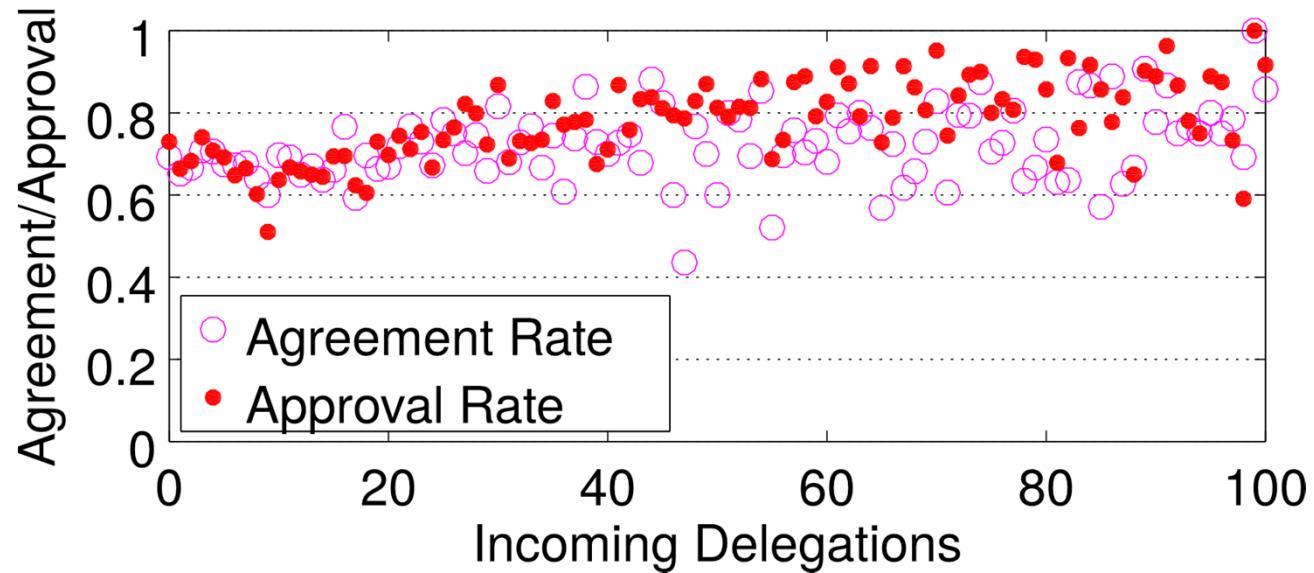


Average Approval Rate



(How many users agree with x% of all voted proposals?)

Average Approval Rate



Powerful voters tend to vote positive
and to agree with the majority

Power

Potential Power:

Measured power in the dataset

Beta power index:

Beta distributed approval rate for Banzhaf index

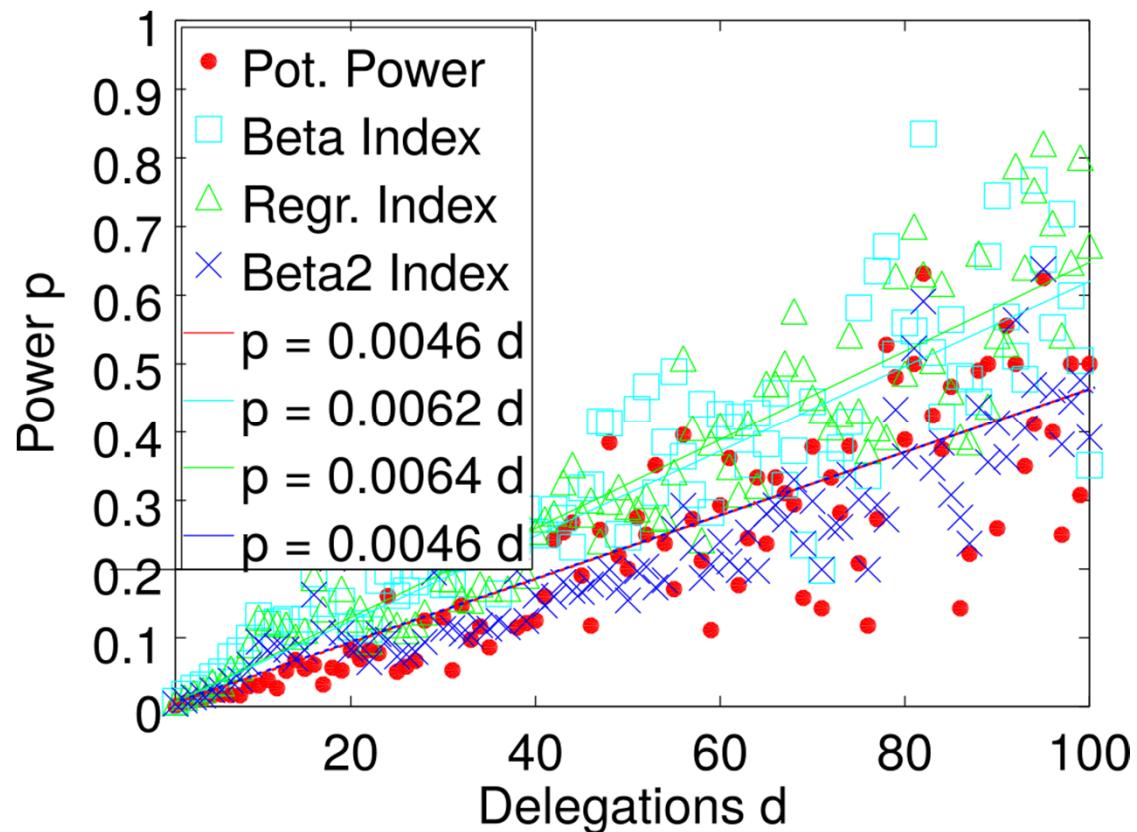
Regression power index:

Logistic regression for predicting the approval rate – given the voting weight – of the Banzhaf index

Beta2 power index:

Beta distributed approval rate for Shapley index

Novel Power Indices



Novel Power Indices

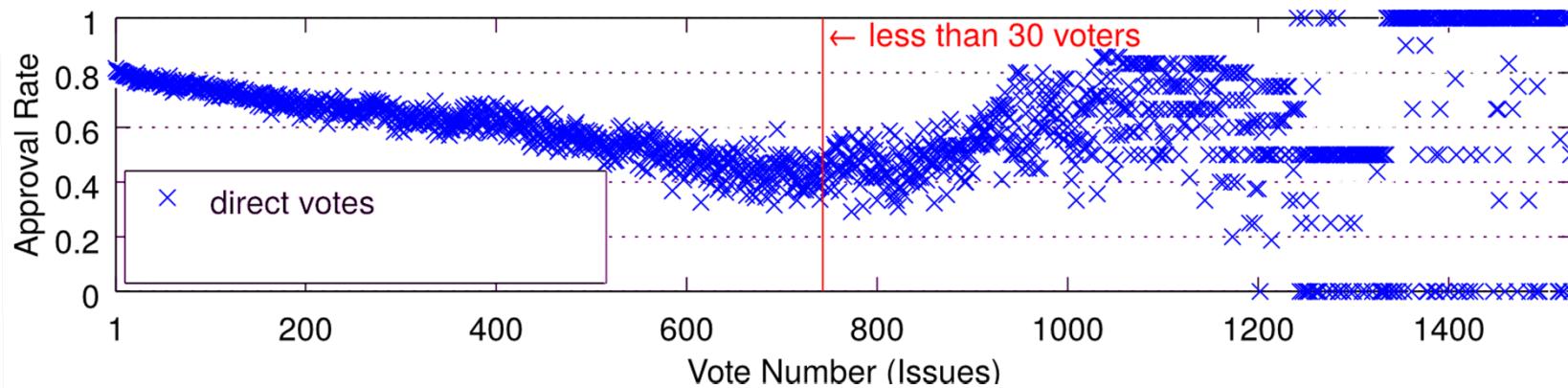
Model	Squared Error	Perplexity
Shapley power index	0.903	78.6
Banzhaf power index	1.320	297.9
Beta power index	2.220	227.8
Regression power index	2.266	232.0
Beta2 power index	0.627	76.6

(Perplexity ~ normalised log likelihood)



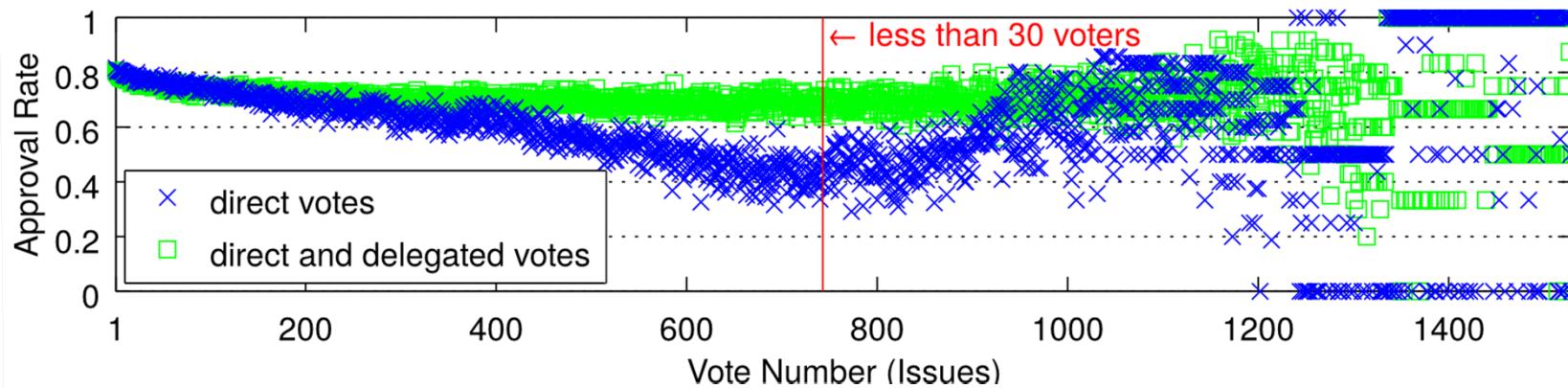
The Impact of Delegations

Approval Rate



→ Approval rate decreases with voting experience

Approval Rate



→ Delegates stabilise the approval rate!

LiquidFeedback der Piratenpartei

[Gastzugang](#) · [Neu registrieren](#) · [Anmelden](#) · [Abstimmen](#)

Hinweise & Neuigkeiten

- **20.05.2015:** Die Domain lqfb.piratenpartei.de wird auf Anweisung des Bundesvorstands in Kürze ins Vorstandspotal umgeleitet werden. Diese Seite ist dann nicht mehr erreichbar.
- **18.05.2015:** Wie im Vorstandspotal [angekündigt](#) sind wir 'angehalten, dieses System erst mal offline zu nehmen'. Wir nutzen diese Gelegenheit, um das System und die Datenbank zu warten. Der Server wird also heute Nacht zweitweise komplett offline sein. Zur Einordnung der Lage verweisen wir im Übrigen auf die aktuelle Nutzerzahlen. Von den run 10.000 registrierten Benutzern haben sich in den letzten Monaten 172 mindestens einmal im System angemeldet. Es existieren aktuell keine neuen oder offenen Initiativen im System.
- **04.10.2013:** Wie im [Vorstandsportal](#) beschrieben, wurden alle von der Antragsfrist des BPT13.2 betroffenen Initiativen heute am 4. Oktober um 23:42 Uhr vorzeitig eingefroren. Die Abstimmung startet automatisch am 11. Oktober und läuft zwei Wochen. Folgende Themen sind betroffen: [3508](#), [3517](#), [3520](#), [3522](#), [3523](#), [3526](#), [3535](#), [3539](#), [3541](#), [3542](#), [3545](#).
- **30.08.2013:** LiquidFeedback läuft jetzt in der Version 2.2.5 (Kern und Frontend). Bisher war es so, dass euer delegiertes Stimmengewicht automatisch inaktiviert wurde, wenn ihr euch ein halbes Jahr am Stück nicht anmeldetet habt. Bei nächsten Login, wurde die Stimme automatisch wieder für mindestens ein weiteres halbes Jahr aktiviert. Dieser "Stimmengewichtsverfall" wurde oft fälschlicherweise als "Delegationsverfall" bezeichnet. In der Zukunft müsst ihr eure Delegationen alle paar Monate explizit bestätigen. Die genaue Zeitspanne steht noch nicht fest. Wenn wir diese Funktion aktivieren, gibt es hier einen entsprechenden Hinweis. Ihr werdet es dann spätestens an den regelmäßige Hinweise im System bemerken.
- Ältere Einträge finden sich im [Archiv](#).

<http://lqfb.piratenpartei.de/> 17.6.2015

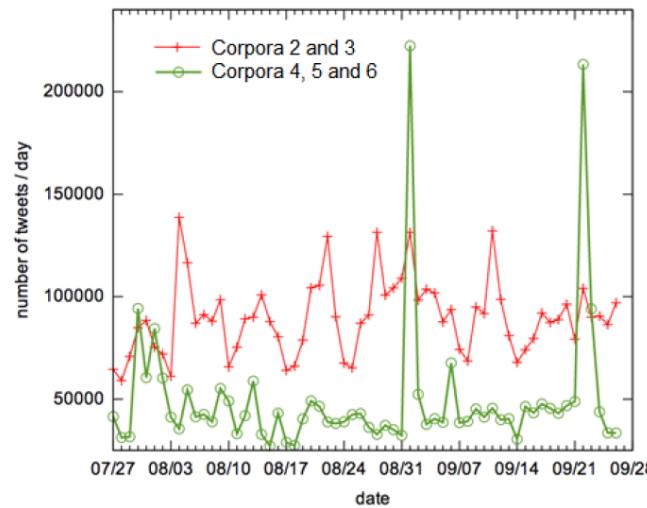
The screenshot shows the homepage of the Piratenpartei Deutschland website. The header features the party's logo (a yellow circle with a black pirate flag) and the text "PIRATENPARTEI Deutschland". Below the header is a navigation menu with links: HOME, POLITIK, PARTEI, MITMACHEN, PRESSE, AKTIONEN, and VIDEO. A sub-navigation bar below the menu includes "Startseite" and the current page title: "Die Abstimmungssoftware LiquidFeedback in der Piratenpartei: wegweisend für Demokratie 2.0? [Update]". The main content area features a large photograph of a party convention where members are raising yellow cards with checkmarks, indicating they have voted using LiquidFeedback. Overlaid on the photo is the article title: "DIE ABSTIMMUNGSSOFTWARE LIQUIDFEEDBACK IN DER PIRATENPARTEI: WEGWEISEND FÜR DEMOKRATIE 2.0? [UPDATE]". At the bottom of the image, the caption reads: "Landesparteitag Piratenpartei Sachsen, Olbernhau | Bild: CC-BY-SA Thomas Grafe". On the left side of the article, there is a small sidebar with the date "31. 05. 2015" and a note: "Ein Guestbeitrag von Christoph Kling* und Jérôme Kunegis**".

The German National Elections 2013 on Twitter

German National Election 2013

Dataset I

	corpus 2 and 3 (candidates and agents)	corpus 4 to 6 (hashtags)
number of tweets	5,573,451	3,088,565
number of handles	356,251	181,927
number of unique hashtags	148,626	168,172



Social Media Monitoring of the Campaigns for the 2013 German Bundestag Elections on Facebook and Twitter Kaczmarek,
Lars; Mayr, Philipp; Vatrapu, Ravi; Bleier, Arним; Blumenberg, Manuela; Gummer, Tobias; Hussain, Abid; Kinder-
Kurlanda, Katharina; Manshaei, Kaveh; Thamm, Mark; Weller, Katrin; Wenz, Alexander; Wolf, Christof

German National Election 2013

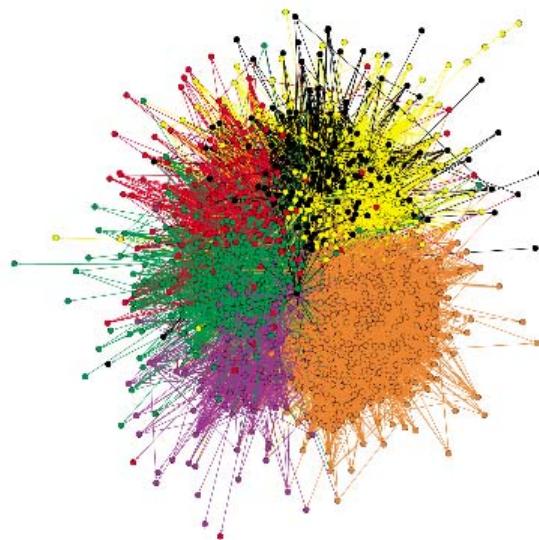
Dataset II

Table 1: Statistics and dataset for the German federal election 2013 on Twitter – parties differ in several interesting ways: Consistently across all conversational practices, the Pirate party exhibits the most homophilic behavior. Mentioning is most strongly used by the Pirates and the Greens ($D = 0.07$), the two largest parties in terms of microblogging politicians (312 and 178), but not in terms of how many votes they actually received (2.2% and 8.4%). D denotes network density, \bar{k} denotes average in/out-degree of nodes, \bar{w} denotes average in/out-weight of nodes, and H denotes homophily, i.e., the tendency of a political party to communicate within party boundaries, computed on the individual level.

Party	Election	Politicians	Following				Retweeting				Mentioning			
	Result		D	\bar{k}_{out}	\bar{k}_{in}	H	D	\bar{w}_{out}	\bar{w}_{in}	H	D	\bar{w}_{out}	\bar{w}_{in}	H
CDU/CSU	41.5%	158	0.15	32	38	0.78	0.08	14	14	0.86	0.04	16	22	0.64
SPD	25.7%	143	0.18	35	41	0.80	0.05	7	9	0.84	0.04	11	17	0.72
FDP	4.8%	143	0.17	35	35	0.78	0.05	7	9	0.84	0.03	6	9	0.55
Greens	8.4%	178	0.21	50	51	0.82	0.08	24	24	0.89	0.07	31	29	0.77
Left	8.6%	97	0.23	30	32	0.79	0.07	8	13	0.91	0.04	13	14	0.68
Pirates	2.2%	312	0.16	57	52	0.89	0.06	40	38	0.93	0.07	73	69	0.92
Total	91.2%	1,031	0.05	44	44	0.83	0.02	25	25	0.90	0.02	39	39	0.79

The German National Elections 2013 on Twitter

Social structure is highly intertwined with semantics

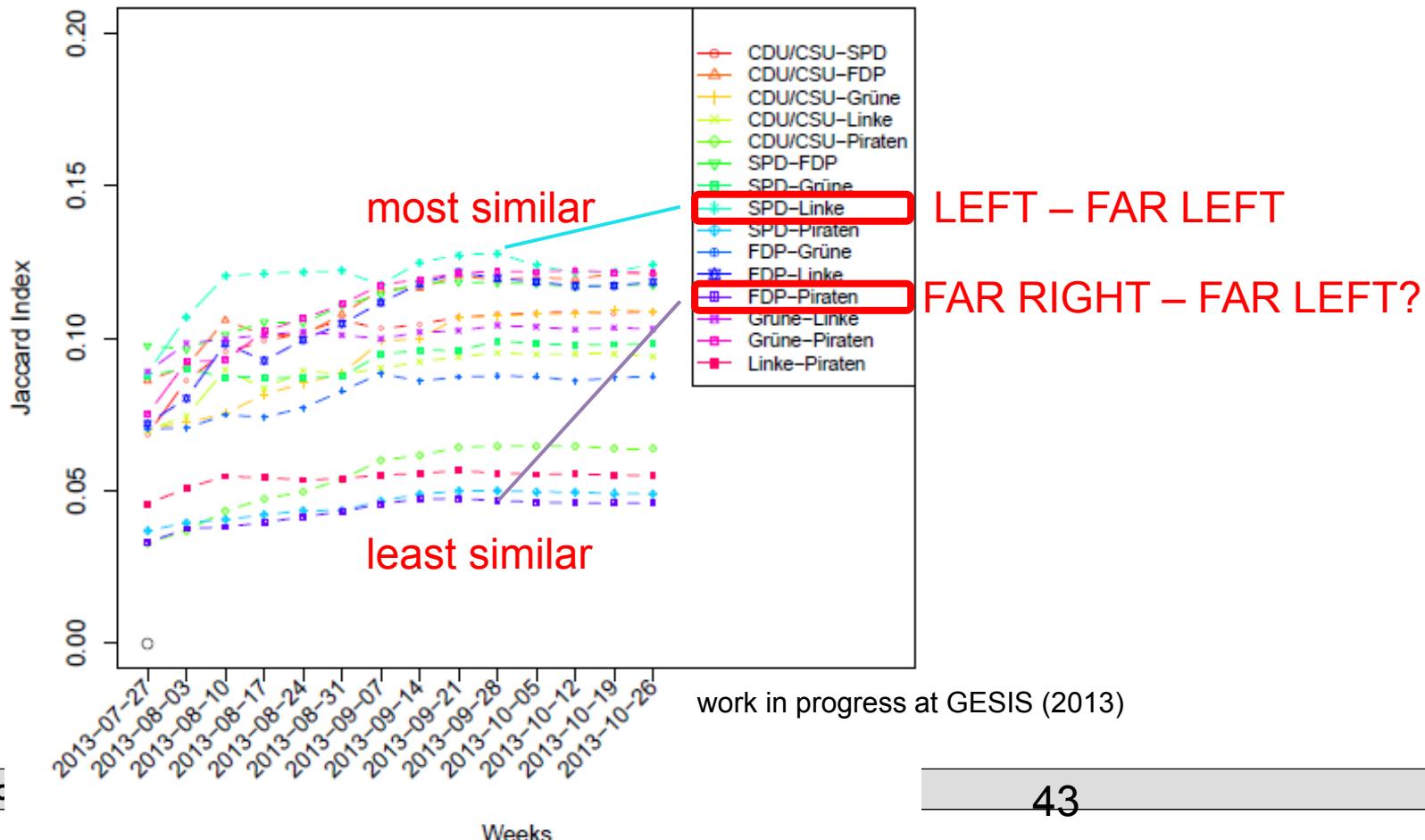


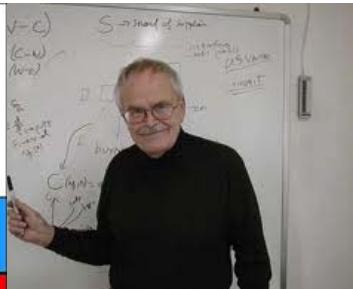
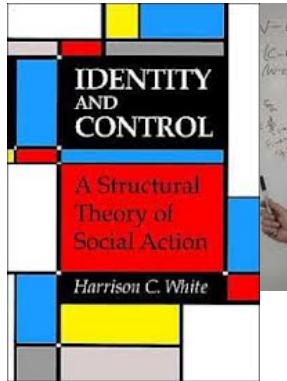
(a) Following ($H=0.83$)

Haiko Lietz, Claudia Wagner, Arnim Bleier, and Markus Strohmaier, When Politicians Talk: Assessing Online Conversational Practices of Political Parties on Twitter, The International AAAI Conference on Weblogs and Social Media (ICWSM2014), Ann Arbor, MI, US, 2014.

German Parliamentary Elections 2013

Similarity between parties as measured by hashtags





Translate theoretical constructs from sociology into quantifiable measure

Theoretical Construct	Measure
Cultural Focus F	Shannon Entropy (Shannon 2001)
Cultural Similarity S	Cosine Similarity (Baeza-Yates and Ribeiro-Neto 1999)
Cultural Reproduction R	Rank Biased Overlap (Wagner et al. 2014)
Institutionness I	Hirsch Index (Hirsch 2005)
Burstiness B	Kleinberg's Burst Weight (Kleinberg 2003)

Following

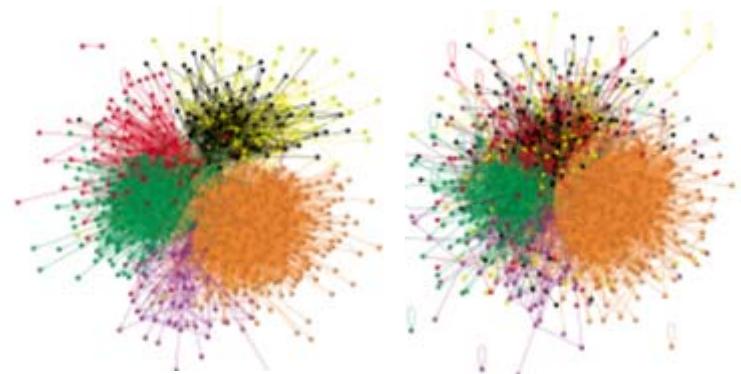


Object of study:
political parties



Unit of analysis:
referenced facts

cataloniawantstovote
abetterbargain
mtholles
standwithourfuture
grüne
obamacareisso
mikromobility
obama
makecollege
steinbrück
p2
btw13
piraten
prism
nsa
iran
cdu
climate
enoughalready
enda
whatwilltake
patrolday
tcof
egypt
snowden
immigration
makecollegeaffordable



Operationalization

Table 2: Operationalization of sociological constructs.

Theoretical Construct	Measure	Description
Cultural Focus F	Shannon Entropy (Shannon 2001)	<i>How strongly does a party focus on cultural facts?</i>
Cultural Similarity S	Cosine Similarity (Baeza-Yates and Ribeiro-Neto 1999)	<i>How similar are parties in terms of their culture vectors?</i>
Cultural Reproduction R	Rank Biased Overlap (Wagner et al. 2014)	<i>How stable is a party's culture vector over time?</i>

German National Election 2013: Political Parties on Twitter

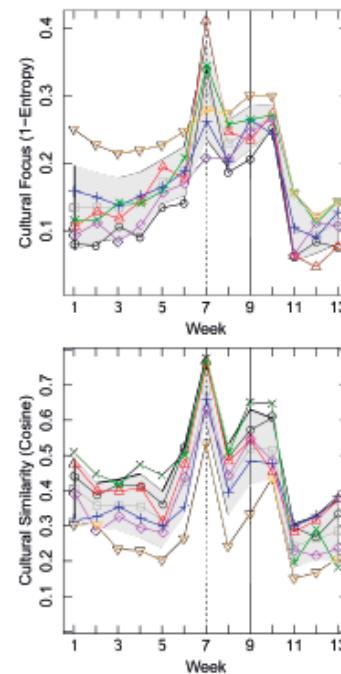


Assessing Cultural focus and Similarity

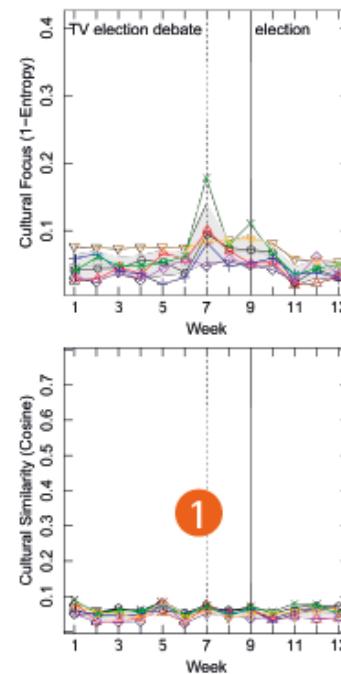
Cultural Focus
How strongly does a party focus on cultural facts?

Cultural Similarity
How similar are parties in terms of their cultural focus?

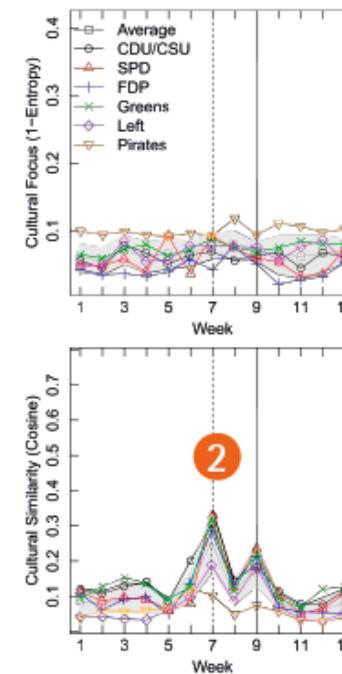
Hashtagging
... to describe communication and find tweets about topics



Retweeting
... forwarding other users' tweets without commenting them



Mentioning
... for the interactive purpose of addressing other users

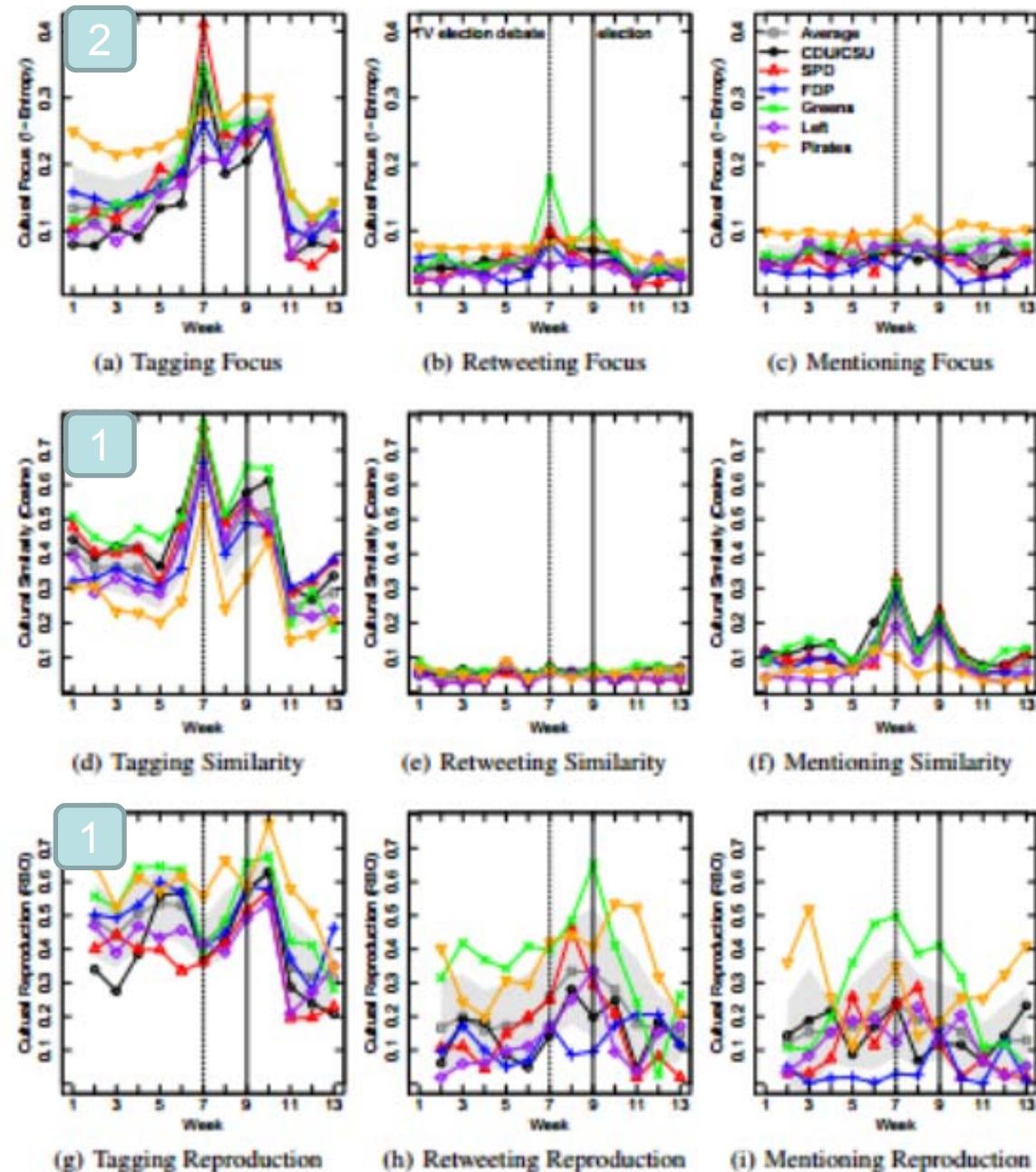


1

Offline events hinder development of a tagging style (i.e., reproduction of unique stable focus)

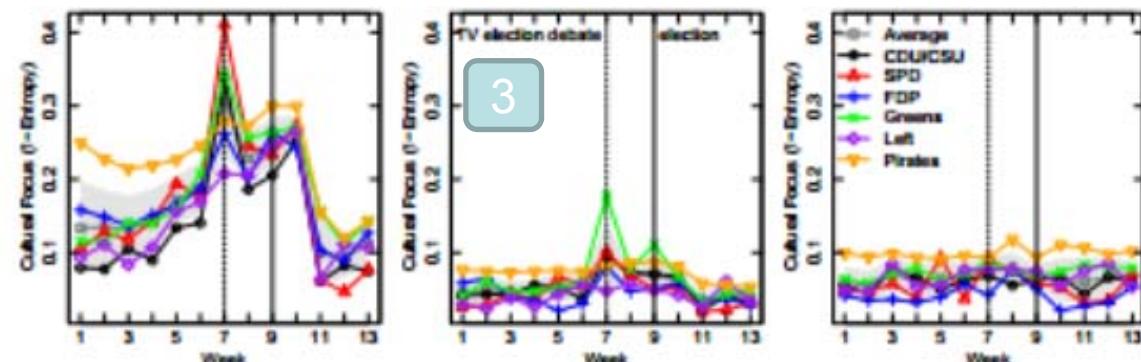
2

Pirates are less impacted by offline events → gain most control → highest homophily in all social networks



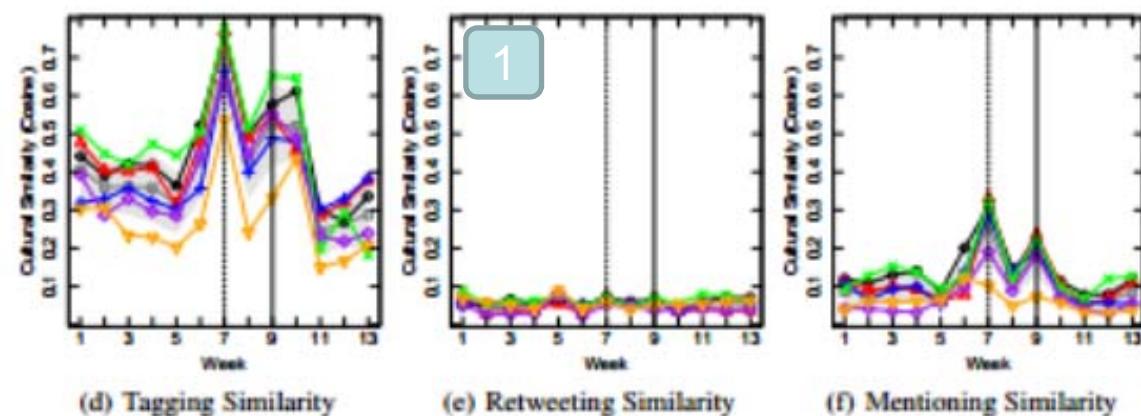
1

Offline events impact retweeting practice less than mentioning and tagging practice. Parties remain unsimilar.



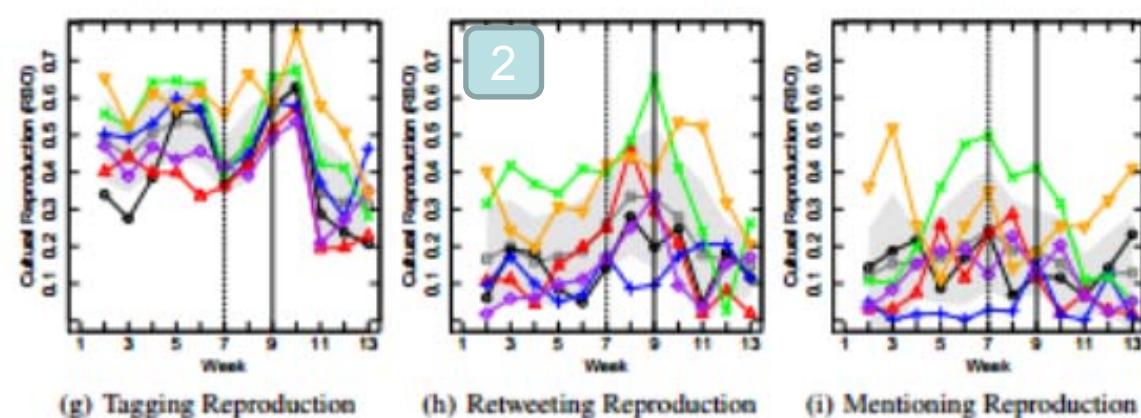
2

All parties reproduce their unique focus, especially Greens and Pirates



3

Greens were impacted by TV Duel → show lower homophily than Pirates

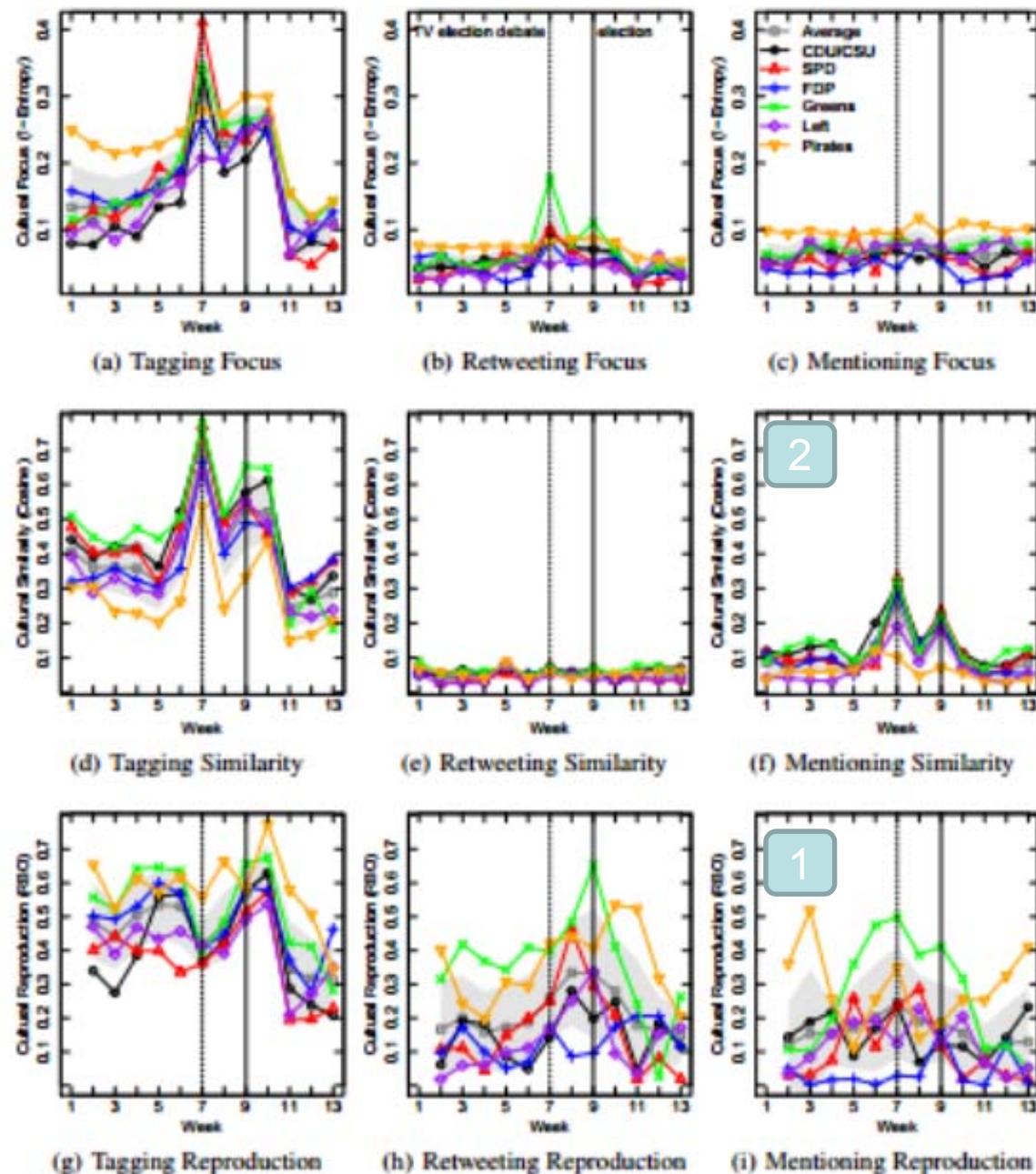


1

No real reproduction trend. Only Greens try to reproduce their unique focus

2

Mentioning practice of all parties (except Pirates) is impacted by offline events



Main Observations

Cross-Party Comparison:

Pirates are the party which has the highest homophily values in all three networks. This suggest that they were most successful in gaining control over their practices. According to White's theory they should have developed a style then.

- Our analysis of the socio-cultural dynamics shows that indeed the Pirate party (and the Greens) reproduced them much more than other parties (see RBO)
- However, Greens were more prone to punctuations than Pirates (see focus and sim)

Cross-practice Comparison:

All parties gained more control over their retweeting practice than their mentioning and tagging practice (see sim)

Punctuations disturb the effort of parties to gain control and develop a style.

- Pirates are the only party which is not effected by punctuations (see similarity) → therefore they are more successful in reproducing themselves and gain more control → higher homophily in all three types of social structure

Cross-Practice Difference RT versus Mention

Retweet network shows higher homophily than mention network

- That means parties are more successful in gaining control in their retweeting practice
- We would expect to see stronger evidence for the formation of styles in the retweeting practice of parties (i.e., focus which is stably reproduced)
- RBO shows that parties on average develop a retweet-style in the phase before the election but it diminishes afterwards. For mention-style the signal is less visible.

Cross-Practice Differences RT versus Mention

Politicians reproduce their focus of retweeting practices successfully before the election (see RBO) → style is developed

They maintain their dissimilarities with other parties – i.e., they stably reproduce their unique focus

Also in the mentioning practice the development of style visible, but is less pronounced (see RBO)

External events impact disturb the mentioning practice. This leads to pointwise observations of higher cross-party similarities.

Conclusions: perturbations hinder or slow down the control gaining process especially in the mentioning practice. Politicians have more control over retweet practice and this practice can be less disturbed by external events.

Cross-Practice Differences Tagging versus RT+Mention

The impact of punctuations of external events is most visible in the tagging practice

- Politicians react on punctuations by adapting their tagging practice. They shift their tagging focus into the same direction → all parties become more similar and diminish their reproduction of cultural facts → therefore they hinder or slow down their development of a style
- After the election
 - it looks like parties stop trying to control their tagging practice; they reproduce their cultural facts less and less, become as unsimilar to other parties as usually and loose their focus.
 - Also for RT practice we see this switch from pre-election style to post election style.
 - For mentioning is less visible since (i) they did not manage to develop a clear style in the pre-election phase and (ii) politicians start the debate about who gets which position and who has to leave now.

Next week

Opinion formation / dynamics and language games.

How does a group of people reach a global consensus
on a given issue?

The naming game

Thank you!