Supplementary materials for article
"Of echo chambers and contrarian clubs:
Exposure to political disagreement among German and Italian users of Twitter"
published on *Social Media + Society*

Cristian Vaccari (Royal Holloway, University of London and University of Bologna)
Augusto Valeriani (University of Bologna)
Pablo Barberà (University of Southern California)
John T. Jost (New York University)
Jonathan Nagler (New York University)
Joshua A. Tucker (New York University)

Published online July 2016

List of Appendices:

Appendix 1 – Keywords employed to retrieve election-related tweets

Germany

Political parties

AFD

CDU

CSU

FDP

Grünen

Linke

Piratenpartei

SPD

Party leaders

Brüderle

Göring-Eckardt

Kipping

Lucke

Merkel

Riexinger

Rösler

Schlömer

Seehofer

Steinbrück

Trittin

Election hashtags

#btw2013

#Bundestagswahl2013

#wahl2013

Italy

Political parties

IDV

Lega

M5S

PD

PDL

Rivoluzione Civile

Scelta Civica

SEL

UDC

Party leaders

Berlusconi

Bersani

Casini

Di Pietro

Grillo

Ingroia

Maroni

Monti

Vendola

Election hashtags

#elezioni2013

Appendix 2 – Assessment of Sample Representativeness

In this section we illustrate evidence suggesting that our respondents can be considered representative of Germans and Italians who discussed the 2013 elections on Twitter, and that the differences that could be measured between these two groups were taken into account in our analysis.

To substantiate this claim, we need to combine and compare information about our respondents and all the users we invited to take our surveys, and to achieve this goal we integrate our survey data with observations of Twitter activity. With regard to our respondents, we know their answers to our survey questions, but – at least for the majority of them – we do not know their social media activities (such as how often they posted messages, how many accounts they followed, and the like) because selection into our surveys was anonymous. With regard to all the users we invited to participate in our surveys, we know their social media activities because Twitter usage is public and we can measure this behavior regardless of whether or not someone chose to take our surveys, but we of course do not know how they would have answered our questions. However, approximately 40% of our respondents chose to provide us with their Twitter handle, so for this sub-sample we have both of these types of information: the answers to our survey questions *and* their behavior on Twitter. Thus we can compare how similar this portion of our respondents are to all the users we invited to participate in our survey (based on Twitter activity) and how similar they are to respondents who did not give us their Twitter handles (via the survey data). We also employ a number of techniques developed by scholars to estimate characteristics of Twitter users such as gender, location, and ideology.

In particular, gender was estimated using a Naive Bayes classifier (Bird, Klein, & Loper, 2009) trained with a list of common Italian and German names and their gender,

and then applied to the name of each Twitter user, as reported on their profile. We found that this technique is able to accurately classify the gender of 90% of Twitter users. Each user's location was identified by parsing the "location" field in each profile using the Data Science Toolkit geocoder (www.datasciencetoolkit.org), which turns text into a set of coordinates, which we then matched to Italian and German regions. We were able to identify the region in which each user lives in 60% of cases in Italy and 51% of cases in Germany. Finally, ideology was measured using the "spatial following model" described in Barberá (2014), which estimates ideology based on the political actors and media outlets that each Twitter user follows. We found that this method is able to classify the self-reported ideological positions (left-right) of Italian respondents in our survey with 82% accuracy and of German respondents with 71% accuracy.

This information in turn allows us to compare Twitter users who answered the surveys with all those we contacted not just in terms of Twitter behavior, but also in terms of socio-demographic and political characteristics, and therefore evaluate whether the former is representative of the latter. As those invited to take the surveys were selected randomly, this process should allow us to ensure that our samples are representative of Germans and Italians who tweeted about the elections and, to the extent that they are not, to reweight the data accordingly.

As shown in Appendix 3, respondents who provided their Twitter usernames were similar to those who answered the survey but did not provide their usernames in terms of demographic characteristics, interest in politics, and ideology. When compared to all the users we asked to participate in the survey, survey respondents turned out to follow more politicians' accounts and to have posted more tweets about the election, although there was no difference in terms of the popularity of different keywords posted by the two groups; in Germany, they were also more likely to be male.

To better ensure that our survey respondents are representative of the populations of Twitter users who tweeted about these elections, we weight our analyses by gender, region, number of political accounts followed, and number of tweets posted that mentioned any of our keywords. This approach is commonly adopted by survey researchers to ensure that sample margins match population margins in a set of key variables (Gelman and Hill 2007, 310-319). For those respondents who did not provide their Twitter usernames, we imputed five sets of values for the latter two variables using a Markov chain Monte Carlo method (Gelman and Hill 2007). We then computed five different sets of weights and ran multiple analyses using each of them, the results of which were then aggregated. Because we only weighted those cases for which we had information concerning all four variables, the total number of cases in our analyses is 999 for Germany and 1,408 for Italy. Therefore, while we do not argue that our samples are representative of the German and Italian populations as a whole, we do feel confident – when using the weighted analyses – that they are fairly representative of Germans and Italians who talked about the 2013 elections on Twitter, at least in so far as their online political activity, gender, and region is concerned.

Appendix 3 – Characteristics of Twitter users invited to take the survey, of Twitter users who participated in the survey and provided their Twitter account names, and of Twitter users who participated in the survey and did not provide their Twitter account names

	Germany			Italy		
Socio-political characteristics	**Did not give username**	**Gave username**	**p-value**	**Did not give username**	**Gave username**	**p-value**
% female	30.5	22.2	.003	40.5	37.3	.210
Average age (years)	34.9	35.1	.841	32.0	32.0	.942
Average educational level (0-1)	.65	.62	.010	.64	.62	.077
Average income bracket (0-1)	.49	.46	.073	.46	.45	.359
Average interest in politics (0-1)	.80	.82	.082	.75	.79	<.001
Ideology (left-right, 0-1)	.32	.33	.537	.38	.37	.803
TOTAL	665	478		880	613	
Twitter activities	**Invited to survey**	**Gave username**	**p-value**	**Invited to survey**	**Gave username**	**p-value**
% female (estimated)	51.7	38.4	<.001	38.2	38.3	.992
Number of followers	865	651	.069	236	239	.883
Total number of tweets	8,695	7,291	.053	3,223	2,983	.574
Tweets mentioning political keywords	6.86	25.04	<.001	12	31	<.001
Number of days since account created	672	930	<.001	605	781	<.001
Number of political accounts followed	5.95	17.49	<.001	11	20	<.001
Ideology (left-right, -3 to +3)[1]	-0.01	-0.03	.239	-.31	-.44	.022
TOTAL	42,647	393		55,245	585	

[1] Ideology was estimated based on the method described in Barberá (2014).

Appendix 4 – Dependent Variable(s): Types of Political Networks Respondents Engage with on Social Media (Respondents Engaging in Neutral Networks Excluded from Analysis)

	Germany			Italy		
	Supportive	**Opposition**	**Mixed**	**Supportive**	**Opposition**	**Mixed**
Face-to-face political networks (ref. Mixed)						
Supportive	1.457**	-1.085	-1.191*	1.556***	-.809*	-1.568***
Oppositional	-.124	.707	-.745	.169	.988**	-1.931***
Political messages exchanged on SNS	.115	-.242*	.130	.068	-.173**	.162
Preferred use of social media	1.751*	-2.095*	-.576	1.852***	-1.806**	-.841
Source of political information						
Internet	.399	.480	-1.170	.568	-.308	-.707
Newspapers	-.796	1.078	.116	-.567	1.088*	-.682
Radio	-1.153*	1.521*	-.056	-.327	-.027	.826
Television	-.599	.129	.667	-.354	-.054	.880
Political efficacy (disagreement with following sentences)						
"People don't have any say"	.570	-.106	-.924	-.425	.420	.231
"Public officials don't care"	-.424	-.296	1.271	.864*	-.505	-.929
"Politics is too complicated"	-.216	-.504	1.222	.089	-.146	.073
Interest in politics	-1.482	1.916	-.183	-1.073	.563	1.378
Trust in political parties	-.016	-1.181	1.275	.280	-.041	-.523
Offline political discussion	.793	-.641	-.629	1.158	-.604	-1.478
Gender (male)	.104	-.539	.508	-.098	.027	.149
Age	.289	-.429	-.381	.277	.423	-1.518
Education	-.851	1.107	.065	.080	-.689	1.111
Income	-1.230	1.165	.977	-.525	.654	-.099
Constant	1.656	-2.381*	-2.654*	-.715	-.308	-1.556
N	322	322	322	560	560	560
Nagelkerke R^2	.291	.327	.176	.250	.281	.159
Log-likelihood	328.511	247.812	211.579	618.599	522.760	348.143

Note: cell entries are log odds (***p≤.001 **p≤.01 *p≤.05).
Dummy variable identifying missing observations for income omitted from table.
All variables range from 0-1 apart from political messages exchanged (0-10) and preferred use of social media (-1 to 1).

Appendix 5 – German Respondents. Dependent Variable: Engagement with supportive political networks on Social Media

	B	s.e.	p	exp(B)	95% CI for exp(B)		99% CI for exp(B)	
					min	max	min	max
Face-to-face political networks (ref. Neutral)								
Supportive	.818	.241	.002	2.265	1.390	3.692	1.178	4.357
Oppositional	-.381	.306	.217	.683	.371	1.258	.304	1.537
Mixed	-.365	.372	.328	.694	.332	1.449	.262	1.837
Political messages exchanged on SNS	.153	.042	.000	1.165	1.073	1.266	1.045	1.299
Preferred use of social media	1.240	.436	.006	3.455	1.447	8.254	1.087	10.983
Source of political information								
Internet	.078	.441	.859	1.082	.455	2.573	.346	3.385
Newspapers	-.460	.330	.164	.631	.330	1.207	.269	1.481
Radio	-1.046	.284	.000	.351	.201	.613	.169	.731
Television	-.406	.332	.223	.666	.345	1.286	.279	1.590
"People don't have any say"	.108	.476	.823	1.114	.413	3.003	.289	4.301
"Public officials don't care"	.167	.437	.704	1.181	.494	2.824	.371	3.760
"Politics is too complicated"	-.660	.348	.058	.517	.261	1.024	.211	1.270
Interest in politics	1.601	.538	.003	4.958	1.722	14.277	1.232	19.948
Trust in political parties	-.276	.504	.587	.759	.274	2.104	.194	2.971
Offline political discussion	.079	.396	.841	1.083	.495	2.366	.386	3.039
Gender (male)	-.128	.208	.541	.880	.583	1.328	.511	1.516
Age	-.712	.553	.200	.491	.165	1.464	.116	2.079
Education	-.431	.503	.393	.650	.239	1.765	.173	2.444
Income	-.582	.437	.183	.559	.237	1.315	.181	1.721
Constant	-.667	.523	.202	.513	.184	1.429	.133	1.973
N	727							
Nagelkerke R^2	.241							
Log-likelihood	834.045							
% explained	69.9							

Dummy variable identifying missing observations for income omitted from table. All variables range from 0-1 apart from political message exchange (0-10) and preferred use of social media (-1 to 1).

Appendix 6 – German Respondents. Dependent Variable: Engagement with oppositional political networks on Social Media

	B	s.e.	p	exp(B)	95% CI for exp(B)		99% CI for exp(B)	
					min	max	min	max
Face-to-face political networks (ref. Neutral)								
Supportive	-.201	.343	.561	.818	.413	1.622	.330	2.029
Oppositional	.899	.350	.012	2.458	1.227	4.924	.979	6.170
Mixed	.489	.492	.322	1.630	.615	4.321	.449	5.921
Political messages exchanged on SNS	-.041	.062	.509	.960	.849	1.085	.816	1.128
Preferred use of social media	-1.032	.578	.078	.356	.113	1.127	.077	1.640
Source of political information								
Internet	1.023	.651	.117	2.782	.774	10.000	.516	14.991
Newspapers	.360	.429	.401	1.434	.619	3.322	.475	4.327
Radio	.612	.436	.165	1.844	.773	4.403	.580	5.860
Television	-.264	.464	.571	.768	.305	1.934	.226	2.614
"People don't have any say"	-.201	.556	.718	.818	.273	2.450	.192	3.477
"Public officials don't care"	-.375	.618	.546	.688	.200	2.359	.134	3.536
"Politics is too complicated"	-.039	.521	.940	.962	.338	2.734	.239	3.866
Interest in politics	-.923	.923	.331	.398	.057	2.770	.028	5.697
Trust in political parties	-.641	.749	.399	.527	.114	2.437	.067	4.154
Offline political discussion	.661	.560	.239	1.937	.641	5.853	.450	8.336
Gender (male)	-.227	.298	.449	.797	.440	1.444	.362	1.755
Age	-1.051	.992	.303	.350	.044	2.792	.020	5.987
Education	1.025	.714	.156	2.787	.670	11.599	.419	18.522
Income	1.421	.797	.086	4.141	.804	21.324	.452	37.950
Constant	-3.087	.945	.003	.046	.007	.317	.003	.625
N	727							
Nagelkerke R^2	.125							
Log-likelihood	518.951							
% explained	86.5							

Dummy variable identifying missing observations for income omitted from table. All variables range from 0-1 apart from political message exchange (0-10) and preferred use of social media (-1 to 1).

Appendix 7 – German Respondents. Dependent Variable: Engagement with mixed political networks on Social Media

	B	s.e.	p	exp(B)	95% CI for exp(B)		99% CI for exp(B)	
					min	max	min	max
Face-to-face political networks (ref. Neutral)								
Supportive	.832	.554	.140	2.297	.755	6.988	.521	10.127
Oppositional	1.200	.553	.031	3.320	1.115	9.879	.788	13.986
Mixed	1.696	.615	.007	5.452	1.606	18.513	1.080	27.531
Political messages exchanged on SNS	.149	.095	.126	1.160	.957	1.406	.898	1.500
Preferred use of social media	.224	.777	.775	1.251	.260	6.021	.153	10.246
Source of political information								
Internet	-.386	.845	.648	.680	.130	3.568	.077	6.016
Newspapers	-.167	.602	.782	.846	.259	2.764	.178	4.021
Radio	.562	.589	.343	1.755	.542	5.685	.368	8.364
Television	-.008	.658	.990	.992	.263	3.748	.167	5.876
“People don't have any say”	-.581	.767	.450	.559	.122	2.554	.075	4.168
“Public officials don’t care”	.849	.799	.290	2.338	.480	11.385	.288	18.959
“Politics is too complicated”	.799	.740	.285	2.223	.506	9.756	.311	15.876
Interest in politics	.846	1.167	.469	2.331	.235	23.107	.114	47.729
Trust in political parties	.597	.912	.517	1.817	.283	11.667	.149	22.162
Offline political discussion	-.118	.800	.883	.889	.184	4.304	.111	7.113
Gender (male)	.367	.458	.427	1.444	.575	3.626	.422	4.936
Age	-.401	1.050	.703	.670	.083	5.431	.042	10.784
Education	-1.042	.905	.252	.353	.059	2.124	.033	3.792
Income	.812	.895	.366	2.253	.382	13.292	.215	23.578
Constant	-5.569	1.221	.000	.004	.000	.043	.000	.093
N	727							
Nagelkerke R^2	.141							
Log-likelihood	312.860							
% explained	93.4							

Dummy variable identifying missing observations for income omitted from table. All variables range from 0-1 apart from political message exchange (0-10) and preferred use of social media (-1 to 1).

Appendix 8 – Italian Respondents. Dependent Variable: Engagement with supportive political networks on Social Media

	B	s.e.	p	exp(B)	95% CI for exp(B)		99% CI for exp(B)	
					min	max	min	max
Face-to-face political networks (ref. Neutral)								
Supportive	.998	.167	.000	2.714	1.956	3.766	1.764	4.175
Oppositional	.040	.192	.836	1.040	.713	1.517	.633	1.710
Mixed	.032	.271	.905	1.033	.607	1.757	.513	2.078
Political messages exchanged on SNS	.120	.038	.003	1.128	1.044	1.217	1.019	1.248
Preferred use of social media	1.714	.355	.000	5.550	2.763	11.147	2.217	13.893
Source of political information								
Internet	.171	.417	.682	1.187	.519	2.716	.397	3.550
Newspapers	-.690	.278	.014	.501	.289	.868	.243	1.036
Radio	-.508	.230	.028	.601	.382	.946	.331	1.092
Television	.078	.258	.762	1.081	.652	1.792	.557	2.100
"People don't have any say"	.042	.231	.855	1.043	.663	1.640	.575	1.890
"Public officials don't care"	.338	.262	.197	1.402	.839	2.343	.714	2.753
"Politics is too complicated"	-.098	.202	.626	.906	.610	1.346	.539	1.524
Interest in politics	-.141	.458	.759	.868	.347	2.170	.257	2.934
Trust in political parties	.379	.301	.209	1.461	.808	2.639	.671	3.180
Offline political discussion	.257	.536	.632	1.293	.452	3.696	.325	5.142
Gender (male)	-.211	.149	.158	.810	.605	1.085	.551	1.190
Age	-.561	.438	.201	.570	.240	1.353	.183	1.782
Education	.361	.492	.463	1.435	.546	3.773	.402	5.123
Income	-.094	.314	.766	.910	.485	1.708	.394	2.104
Constant	-1.348	.506	.008	.260	.096	.700	.071	.955
N	1167							
Nagelkerke R^2	.164							
Log-likelihood	1391.319							
% explained	68.7							

Dummy variable identifying missing observations for income omitted from table. All variables range from 0-1 apart from political message exchange (0-10) and preferred use of social media (-1 to 1).

Appendix 9 – Italian Respondents. Dependent Variable: Engagement with oppositional political networks on Social Media

	B	s.e.	p	exp(B)	95% CI for exp(B)		99% CI for exp(B)	
					min	max	min	max
Face-to-face political networks (ref. No opinion)								
Supportive	.183	.296	.538	1.201	.668	2.157	.554	2.604
Oppositional	1.756	.251	.000	5.788	3.532	9.485	3.020	11.094
Mixed	1.082	.348	.002	2.950	1.489	5.846	1.200	7.254
Political messages exchanged on SNS	-.058	.046	.211	.944	.862	1.033	.838	1.063
Preferred use of social media	-1.056	.462	.022	.348	.141	.861	.106	1.146
Source of political information								
Internet	-.488	.509	.339	.614	.225	1.675	.163	2.307
Newspapers	.060	.369	.872	1.062	.512	2.200	.406	2.779
Radio	.005	.308	.987	1.005	.549	1.839	.454	2.227
Television	-.068	.359	.850	.934	.460	1.897	.367	2.376
“People don't have any say”	.469	.311	.132	1.598	.868	2.942	.716	3.568
“Public officials don’t care”	-.239	.395	.546	.787	.360	1.723	.279	2.219
“Politics is too complicated”	.000	.266	1.000	1.000	.594	1.683	.504	1.982
Interest in politics	.805	.556	.148	2.238	.751	6.671	.531	9.422
Trust in political parties	-.386	.412	.349	.680	.303	1.526	.235	1.970
Offline political discussion	-1.014	.692	.144	.363	.093	1.413	.061	2.170
Gender (male)	.041	.201	.838	1.042	.703	1.546	.621	1.750
Age	-.347	.552	.530	.707	.239	2.086	.170	2.931
Education	.290	.617	.639	1.336	.398	4.480	.272	6.552
Income	.227	.368	.537	1.255	.610	2.586	.485	3.247
Constant	-1.893	.660	.004	.151	.041	.549	.027	.825
N	1167							
Nagelkerke R^2	.157							
Log-likelihood	889.100							
% explained	85.1							

Dummy variable identifying missing observations for income omitted from table. All variables range from 0-1 apart from political message exchange (0-10) and preferred use of social media (-1 to 1).

Appendix 10 – Italian Respondents. Dependent Variable: Engagement with mixed political networks on Social Media

	B	s.e.	p	exp(B)	95% CI for exp(B)		99% CI for exp(B)	
					min	max	min	max
Face-to-face political networks (ref. No opinion)								
Supportive	.580	.391	.140	1.786	.825	3.866	.644	4.950
Oppositional	.228	.404	.572	1.256	.569	2.775	.443	3.562
Mixed	2.072	.446	.000	7.942	3.274	19.268	2.456	25.690
Political messages exchanged on SNS	.126	.066	.055	1.135	.997	1.291	.957	1.345
Preferred use of social media	-.675	.680	.322	.509	.134	1.940	.088	2.965
Source of political information								
Internet	-.983	.748	.189	.374	.086	1.621	.054	2.570
Newspapers	-.433	.521	.408	.648	.231	1.823	.165	2.545
Radio	.815	.449	.070	2.259	.936	5.452	.709	7.194
Television	.911	.578	.117	2.487	.793	7.794	.550	11.249
"People don't have any say"	-.048	.427	.911	.953	.412	2.204	.317	2.868
"Public officials don't care"	-.886	.543	.103	.412	.142	1.198	.101	1.678
"Politics is too complicated"	.311	.393	.429	1.365	.631	2.952	.495	3.766
Interest in politics	1.969	1.059	.074	7.164	.811	63.246	.376	136.337
Trust in political parties	-.768	.605	.205	.464	.142	1.521	.097	2.212
Offline political discussion	-.587	1.384	.675	.556	.033	9.400	.012	25.078
Gender (male)	.212	.287	.460	1.236	.705	2.168	.590	2.588
Age	-.798	.833	.340	.450	.087	2.334	.051	3.957
Education	.614	.964	.525	1.848	.275	12.431	.149	22.930
Income	-.528	.571	.357	.590	.190	1.827	.132	2.633
Constant	-4.822	1.210	.000	.008	.001	.089	.000	.195
N	1167							
Nagelkerke R^2	.143							
Log-likelihood	503.732							
% explained	93.4							

Dummy variable identifying missing observations for income omitted from table. All variables range from 0-1 apart from political message exchange (0-10) and preferred use of social media (-1 to 1).

Appendix 11 – Question wording, response modes and descriptive statistics for all control variables included in the models

Question wording	Response modes	Variable[2]
Overall, how important are social media to you personally when it comes to finding other people who share your views about important political issues? Overall, how important are social media to you personally when it comes to finding other people who do not share your views about important political issues?	Very important; somehow important; not too important; Not at all important; (Don't know).	"Preferred use of social media"[3] Italian dataset *Mean .017; Standard Deviation .201* German dataset *Mean .039; Standard Deviation .240*
Generally speaking, how much are you interested in politics?	Very interested; Moderately interested; Slightly interested; Not interested at all; (Don't know).	"Interest in politics" Italian dataset *Mean .773; Standard Deviation .227* German dataset *Mean .791; Standard Deviation .242*
To what extent do you agree or disagree with each of these statements? People like me don't have any say about what the government does.	Definitely not true; Not really true; Quite true; Definitely true; (Don't Know).	"People don't have any say" (disagreement with) Italian dataset *Mean .483; Standard Deviation .316* German dataset *Mean .540 Standard Deviation .279*
To what extent do you agree or disagree with each of these statements? Public officials care what people like me think.	Definitely not true; Not really true; Quite true; Definitely true; (Don't Know).	"Public officials don't care" (disagreement with) Italian dataset *Mean .341; Standard Deviation .288* German dataset

[2] All variables included in our models range from 0-1 apart from preferred use of social media (-1 to 1). Descriptive statistics presented in the table refer to these variables (weighted as described in the "Improving the representativeness of the sample" sub-section in the article).

[3] This variable was built based on responses to the two questions, which we combined by subtracting the values of the second from those of the first, so that respondents who claimed that agreement was more important than disagreement would have positive values, whereas those who stated that disagreement was more important than agreement would have negative values, and those who attributed equal importance to agreement and disagreement would have a score of zero.

		Mean .416; Standard Deviation .268
To what extent do you agree or disagree with each of these statements? Sometimes, politics and government seem so complicated that a person like me can't really understand what's going on,	Definitely not true; Not really true; Quite true; Definitely true; (Don't Know).	"Politics is too complicated" (disagreement with) Italian dataset *Mean .536; Standard Deviation .346* German dataset *Mean .622; Standard Deviation .297*
How much do you trust the following institutions and organizations? Political parties.	A lot; A fair amount; A little; Not at all; (Don't know).	"Trust in political parties" Italian dataset *Mean .242; Standard Deviation .255* German dataset *Mean .386; Standard Deviation .246*
How often do you talk about politics with your friends, family, and acquaintances?	Every day or almost every day; A few times a week; A few times a month; Never or almost never; (Don't know).	"Offline political discussion (frequency)" Italian dataset *Mean .880; Standard Deviation .179* German dataset *Mean .657; Standard Deviation .316*
How often do you turn to each of the following media outlets for getting political news of your interest? Internet.	Never; At least once a month; At least once a week; Everyday; More than once per day; (Don't know).	"Source of political information: Internet" Italian dataset *Mean .843; Standard Deviation .210* German dataset *Mean .789; Standard Deviation .275*
How often do you turn to each of the following media outlets for getting political news of your interest? Newspapers.	Never; At least once a month; At least once a week; Everyday; More than once per day; (Don't know).	"Source of political information: Newspapers" Italian dataset *Mean .410; Standard Deviation .295* German dataset *Mean .338; Standard Deviation .307*
How often do you turn to each of the following media outlets for getting political news of your interest? Radio.	Never; At least once a month; At least once a week; Everyday; More than once per day; (Don't know).	"Source of political information: Radio" Italian dataset *Mean .528; Standard Deviation .325* German dataset *Mean .496; Standard Deviation .357*

How often do you turn to each of the following media outlets for getting political news of your interest? Television.	Never; At least once a month; At least once a week; Everyday; More than once per day; (Don't know).	"Source of political information: Television" Italian dataset *Mean .633; Standard Deviation .278* German dataset *Mean .488; Standard Deviation .328*
Are you…	Male; Female.	"Gender (male)" Italian dataset *Male 56,6% of respondents* German dataset *Male 65,4% of respondents*
What year were you born in? [Italian dataset] How old are you? [German dataset]	[Open answer] Italian dataset (Youngest resp. aged 13, oldest 76) German dataset (Youngest resp. aged 13, oldest 72)	"Age" Italian dataset *Mean .308; Standard Deviation .182* German dataset *Mean .351; Standard Deviation .207*
What best describes your final level of education?	None; Primary school; Secondary School degree; Vocational school degree; High school degree; Undergraduate degree; Postgraduate degree; Doctoral Degree. [Italian dataset] None; Primary school; Secondary school (Hauptschule); High School (Abitur); Undergraduate degree; Postgraduate degree; Doctoral Degree. [German dataset]	"Education" Italian dataset *Mean .636; Standard Deviation .153* German dataset *Mean .622; Standard Deviation .214*
What is the gross annual income, before tax or other deductions, for you and your family? [Italian dataset] What is the gross monthly income, before tax or other deductions, for you and your family? [German dataset]	Less than 6,000 Euro; Between 6,000 and 12,000 Euro; Between 12,000 and 18,000 Euro; Between 18,000 and 24,000 Euro; Between 24,000 and 30,000 Euro; Between 30,000 and	"Income"[4] Italian dataset *Mean .463; Standard Deviation .250* German dataset *Mean .455; Standard Deviation .231*

[4] Because our income variable included a large proportion of missing data we mean-replaced these missing values and added a dummy variable to the analysis to identify these cases (further explanations of the method and implications are in Note 16 in the article).

	36,000 Euro; Between 36,000 and 42,000 Euro; Between 42,000 and 50,000 Euro; Between 50,000 and 75,000 Euro; More than 100,000 Euro. [Italian dataset] Less than 500 Euro; Between 500 and 900 Euro; Between 900 and 1,300 Euro; Between 1,300 and 1,500 Euro; Between 1,500 and 2,000 Euro; Between 2,000 and 2,600 Euro; Between 2,600 and 3,500 Euro; Between 3,500 and 4,500 Euro; Between 4,500 and 6,000 Euro; Between 6,000 and 8,000 Euro; More than 8,000 Euro. [German dataset]	

Appendix 12 – Dependent Variable(s): Types of Political Networks Respondents Engage with on Social Media (Variable Measuring Preference for Encountering Agreeing vs Disagreeing Viewpoints on Social Media Excluded from Models)

	Germany			Italy		
	Supportive	**Opposition**	**Mixed**	**Supportive**	**Opposition**	**Mixed**
Face-to-face political networks (ref. Neutral)						
Supportive	.890*** (.223)	-.262 (.351)	.831 (.545)	1.040*** (.165)	.198 (.292)	.681 (.384)
Oppositional	-.355 (.283)	1.027*** (.312)	1.286* (.563)	.087 (.193)	1.751*** (.247)	.243 (.404)
Mixed	-.280 (.368)	.392 (.526)	1.720** (.607)	.079 (.266)	1.075** (.345)	2.094*** (.438)
Political messages exchanged on SNS	.133** (.042)	-.038 (.064)	.168 (.095)	.131*** (.037)	-.064 (.045)	.097 (.065)
Source of political information						
Internet	.126 (.449)	.837 (.598)	-.311 (.826)	.152 (.408)	-.480 (.504)	-.882 (.739)
Newspapers	-.522 (.310)	.297 (.419)	-.039 (.589)	-.728** (.267)	.049 (.355)	-.438 (.503)
Radio	-.829** (.273)	.549 (.428)	.477 (.565)	-.534* (.221)	-.006 (.302)	.714 (.435)
Television	-.391 (.303)	-.216 (.458)	.088 (.656)	.117 (.253)	-.014 (.351)	.854 (.567)
Political efficacy (disagreement with following sentences)						
"People don't have any say"	.185 (.426)	-.148 (.546)	-.430 (.728)	.079 (.228)	.441 (.308)	-.164 (.421)
"Public officials don't care"	-.130 (.415)	-.134 (.608)	.748 (.773)	.340 (.258)	-.236 (.393)	-.852 (.531)
"Politics is too complicated"	-.456 (.329)	-.001 (.503)	.842 (.748)	-.048 (.197)	-.005 (.262)	.264 (.379)
Interest in politics	1.412** (.502)	-.644 (.831)	.784 (1.105)	-.288 (.444)	.912 (.551)	2.272* (1.058)
Trust in political parties	.084 (.494)	-.955 (.721)	.534 (.863)	.451 (.298)	-.445 (.405)	-.601 (.585)
Offline political discussion	-.002 (.368)	.758 (.515)	-.112 (.777)	.421 (.526)	-1.095 (.675)	-.523 (1.359)
Gender (male)	-.105 (.202)	-.143 (.289)	.378 (.460)	-.203 (.146)	.029 (.199)	.216 (.282)
Age	-.674 (.533)	-.925 (.891)	-.388 (1.017)	-.487 (.427)	-.381 (.548)	-.890 (.815)
Education	-.434 (.495)	.838 (.685)	-1.089 (.896)	.377 (.486)	.129 (.613)	.555 (.941)
Income	-.558 (.422)	1.159 (.746)	.822 (.849)	-.094 (.307)	.241 (.359)	-.381 (.552)
Constant	-.673 (.478)	-3.179*** (.879)	-5.799*** (1.166)	-1.512** (.499)	-1.776** (.646)	-4.948*** (1.194)
N	776	776	776	1193	1193	1193
Nagelkerke R^2	.205	.113	.145	.137	.148	.138
Log-likelihood	914.299	549.283	323.175	1439.763	907.024	525.302

Note: cell entries are log odds (***p≤.001 **p≤.01 *p≤.05), S.E. in parentheses. Dummy variable identifying missing observations for income omitted from table. All variables range from 0-1 apart from political messages exchanged (0-10).

Appendix 13 – Bivariate relationship between online and offline networks of political discussion, Germany

		Offline Networks				
		Supportive	Mixed	Oppositional	Neutral	N
Online Networks	Supportive	56.4%	33.1%	32.9%	29.6%	369
	Mixed	6.8%	15.8%	8.8%	3%	60
	Oppositional	8.3%	13%	23.8%	11.6%	114
	Neutral	28.5%	38.1%	34.5%	55.8%	366
	N	345	79	154	331	909

Appendix 14 – Bivariate relationship between online and offline networks of political discussion, Italy

		Offline Networks				
		Supportive	Mixed	Oppositional	Neutral	N
Online Networks	Supportive	52.4%	29%	27.9%	25.5%	478
	Mixed	7.4%	20.9%	6.1%	4.3%	97
	Oppositional	9.1%	19.7%	32.1%	8%	207
	Neutral	31.1%	30.4%	33.9%	62.2%	556
	N	465	112	340	421	1338

References

Barberá, P. (2014). Birds of the Same Feather Tweet Together. Bayesian Ideal Point Estimation Using Twitter Data. *Political Analysis*, *23*(1), 76-91.

Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with Python*. Sebastopol, CA: O'Reilly.

Gellman, A., & Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. New York: Cambridge University Press.