Fake It Til You Make It: A Natural Experiment to Identify European Politicians' Benefit from Twitter Bots

Online Appendix

American Political Science Review

Bruno Castanho Silva
University of Cologne
bcsilva@wiso.uni-koeln.de

Sven-Oliver Proksch
University of Cologne
so.proksch@uni-koeln.de

August 11, 2020

Contents

1 Descriptives of Twitter Data

Table A.1: Number of Tweets by Country

Country	All Tweets	Active users
Austria	33081	55
Belgium	24487	131
Bulgaria	276	8
Croatia	3908	25
Cyprus	6400	21
Czechia	8683	46
Denmark	47353	138
Estonia	7709	35
Finland	104520	158
France	286750	438
Germany	189474	422
Greece	66374	119
Hungary	1413	6
Ireland	101805	137
Italy	207517	610
Latvia	15224	53
Lithuania	1814	19
Luxembourg	1673	25
Malta	32559	51
Netherlands	96268	137
Poland	95463	278
Portugal	7574	47
Romania	99	5
Slovakia	950	15
Slovenia	21125	25
Spain	731812	271
Sweden	208234	264
Uk	623873	489

Table A.2: Top 3 Most Popular Accounts by Party Family

Party family	Country	Followers before July 2018	Name
Agrarian/centre	Finland	118523	Juha Sipilä
Agrarian/centre	Sweden	111069	Annie Lööf
Agrarian/centre	Poland	64418	W.Kosiniak-Kamysz
Christian Democrat	Germany	235069	Peter Altmaier
Christian Democrat	Germany	193672	Peter Tauber
Christian Democrat	Ireland	167186	Leo Varadkar
Confessional	Netherlands	76804	Kees van der Staaij
Confessional	Netherlands	34386	Gert-Jan Segers
Confessional	Netherlands	22671	Elbert Dijkgraaf
Conservative	Spain	1668178	Mariano Rajoy Brey
Conservative	UK	876726	Boris Johnson
Conservative	Italy	603957	Vittorio Sgarbi
Green	Netherlands	246166	Jesse Klaver
Green	Germany	130391	K. Göring-Eckardt
Green	Netherlands	126016	Marianne Thieme
Liberal	Spain	1054988	Albert Rivera
Liberal	Netherlands	691325	Alexander Pechtold
Liberal	France	403728	Bruno Le Maire
No family	Italy	405502	Luigi Di Maio
No family	Czechia	375581	Andrej Babis
No family	Italy	229150	Alessandro Di Battista
Radical Left	Spain	2310758	Pablo Iglesias
Radical Left	France	1919100	Jean-Luc Mélenchon
Radical Left	Spain	806964	Íñigo Errejón
Radical Right	France	2132753	Marine Le Pen
Radical Right	Netherlands	955994	Geert Wilders
Radical Right	Poland	373844	Beata Szydło
Regionalist	Spain	589316	Gabriel Rufián
Regionalist	Spain	297829	Joan Tardà i Coma
Regionalist	Ireland	161683	Gerry Adams
Socialist	UK	1836572	Jeremy Corbyn
Socialist	Spain	889975	Pedro Sánchez
Socialist	Italy	860446	Laura Boldrini

2 Changes by Country and Party

Figure A.1: Ten Parties with the Largest Percentage Loss of Followers during the Purge

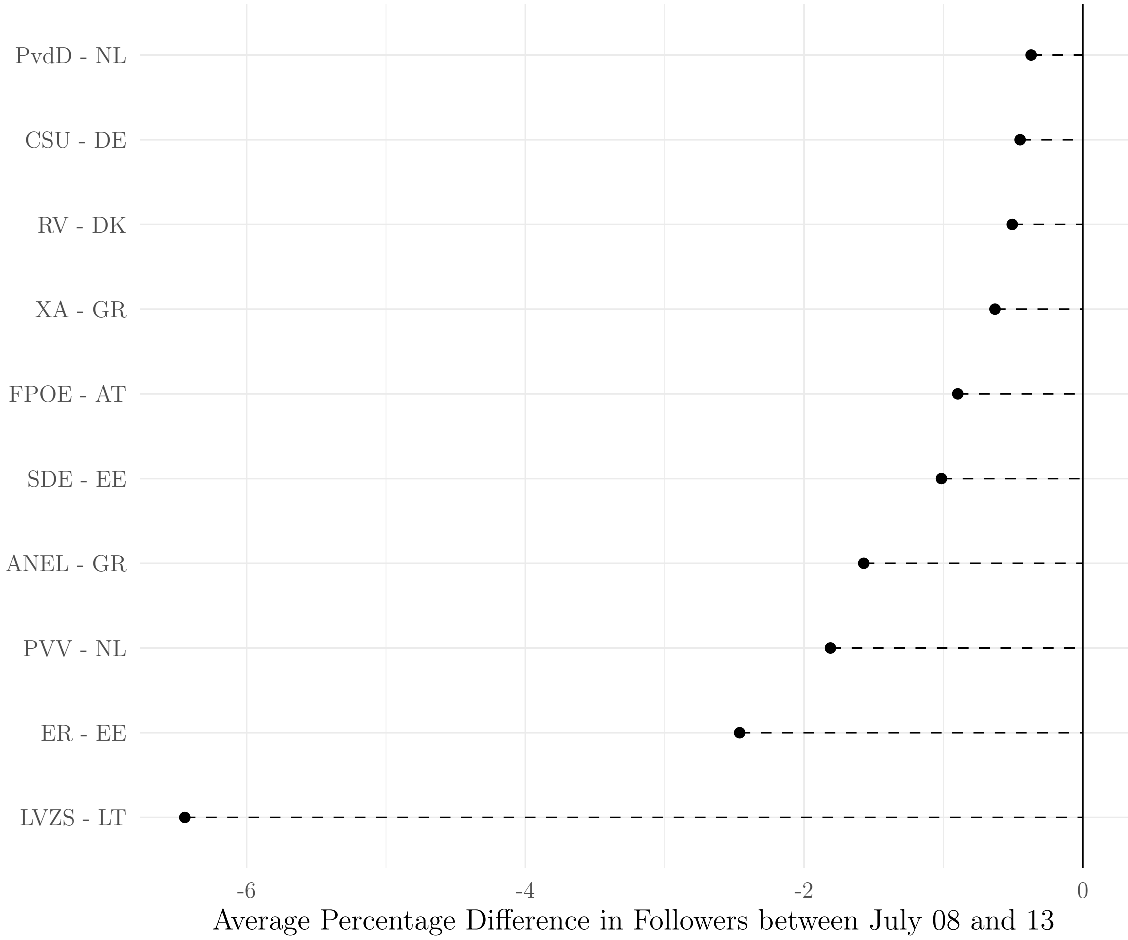
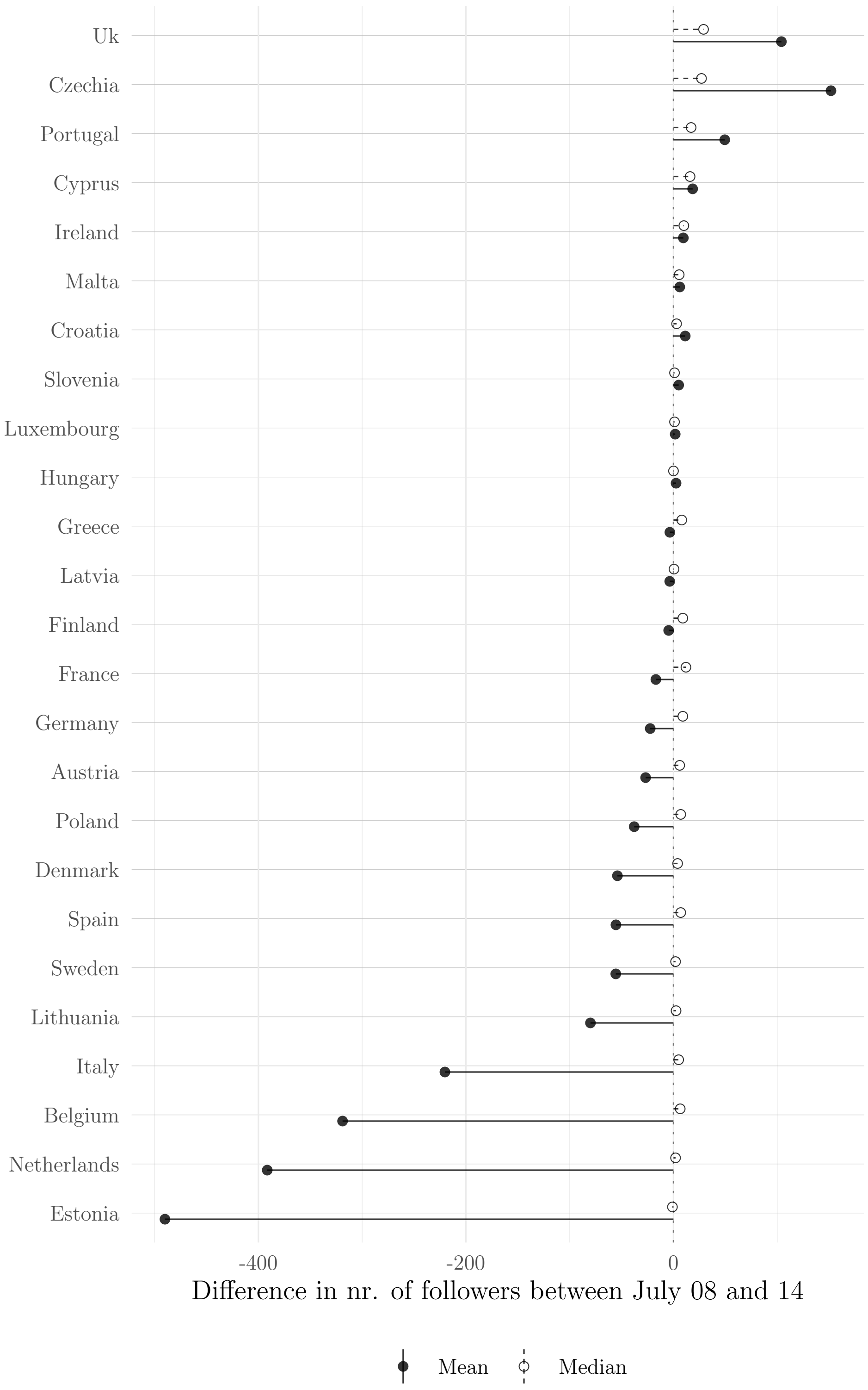


Figure A.2: Change in MPs' Followers by Country



3 Growth rate

The Table below presents the average growth rate for politicians each party family for 100 days before and 100 days after the purge, divided into groups based on how many followers the politicians had at the beginning of each period.

Table A.3: Growth Rate in Number of Followers, 100 days before and after the Purge.

Family	Followers Before July	After	Before
Conservative	Fewer than 10K followers	9.95	13.19
Green	Fewer than 10K followers	8.21	14.73
Liberal	Fewer than 10K followers	14.84	23.36
Radical Left	Fewer than 10K followers	10.53	19.56
Radical Right	Fewer than 10K followers	8.46	11.7
Socialist	Fewer than 10K followers	6.96	9.4
Conservative	From 10K to 50K	8.94	7.93
Green	From 10K to 50K	6.42	9.16
Liberal	From 10K to 50K	10.8	11.5
Radical Left	From 10K to 50K	6.71	10.19
Radical Right	From 10K to 50K	5.11	5.69
Socialist	From 10K to 50K	6.14	8.27
Conservative	More than 50K followers	6.31	7.09
Green	More than 50K followers	2.67	7.13
Liberal	More than 50K followers	3.87	7
Radical Left	More than 50K followers	2.26	5.41
Radical Right	More than 50K followers	2.19	5.54
Socialist	More than 50K followers	3.14	5.84

Notes: average percentage change in the number of followers for politicians in each party family, for the 100 days **After** the purge and 100 days **Before** the purge.

4 EU Filter Terms

Table A.4: List of EU Filter Terms

Language	Stems
Bulgarian	ЕС, евросъюз, европ, брюксел
Croatian	EU, europ, bruxelles, brussel, bruselj
Danish	EU, europ, bryssel, bruxelles
German	EU, europ, brüssel
Greek	ΕΕ, ευρωπ, βρυξέλλες
English	EU, europe , brussels
Spanish	UE, europ , bruselas
Estonian	EL, euroop , brüssel
Finnish	EU, euroop , bryssel
French	UE, europ , bruxelles
Hungarian	EU, europ , brüszel
Italian	UE, europ , bruxelles
Lithuanian	ES, europ , briusel
Latvian	ES, eirop , brisel
Dutch	EU, europ, brussel
Polish	UE, europ, bruksel
Portuguese	UE, europ, bruxelas
Romanian	UE, europ, bruxelles
Slovak	EU, europ, brusel
Slovenian	EU, evrop, bruselj
Swedish	EU, europ, bryssel
Twitter handles (included in all)	eu_commission, europarl_en, eucouncil, junckereu, eucopresident, ep_president, ep_presschulz, @coe, aldeparty, europarlpress, europarl_fr, eurlex, eucourtpress, euauditors, euombudsman, eu_eeas, europarl_it, euatun, jmdbarroso, ecb, eucouncilpress, @epp, eppgroup, theprogressives, pes_pse, aldegroup, guengl, greensep, ecrgroup, addeurope, enf_ep, enl_france, groupeenl, efdgroup

5 Alternative Model Specifications

5.1 Increased time windows

The following models increase the time-window around which tweets are considered to count the number of followers of users. This increases the number of MPs included, since it adds those who might not have tweeted one week before/after July 11, but did so two or four weeks before/after.

Table A.5: Individual and Party-Level Determinants of Percentage Changes in Followers – July 09-13; using tweets from June 24 to July 27

	Model 1	Model 2	Model 3	Model 4
Intercept	4.82*	5.15*	5.22*	5.15*
	[3.64; 6.01]	[3.85; 6.59]	[3.84; 6.69]	[3.75; 6.58]
Male	.00	.00	.01	.00
	[−.24; .25]	[−.24; .26]	[−.24; .25]	[−.24; .25]
Terms in office	−.03	−.03	−.03	−.03
	[−.12; .07]	[−.13; .07]	[−.11; .07]	[−.12; .07]
Cabinet experience	.15*	.15*	.15*	.15*
	[.03; .28]	[.03; .28]	[.02; .28]	[.03; .27]
Twitter Sentiment	−.30	−.31	−.30	−.31
	[−.82; .22]	[−.80; .20]	[−.81; .18]	[−.79; .20]
Twitter EU Sentiment	.05	.06	−.00	.07
	[−.09; .20]	[−.10; .20]	[−.16; .15]	[−.08; .24]
Nr. of Followers (log)	−.50*	−.50*	−.51*	−.50*
	[−.59; −.41]	[−.60; −.41]	[−.60; −.41]	[−.60; −.42]
Seat share	.00	.00	.00	.00
	[−.02; .02]	[−.02; .02]	[−.01; .02]	[−.02; .02]
In government	−.10	−.11	−.08	−.12
	[−.60; .38]	[−.62; .43]	[−.56; .36]	[−.62; .39]
EU Position	−.00	−.06	−.07	−.06
	[−.13; .12]	[−.25; .11]	[−.26; .11]	[−.24; .13]
Radical right		−.65	−.83	−.65
		[−1.83; .38]	[−1.92; .29]	[−1.80; .47]
Radical left		.11	.11	.12
		[−.73; .88]	[−.73; .87]	[−.70; .96]
Twitter EU Sentiment * Radical right			1.38*	
			[.59; 2.11]	
Twitter EU Sentiment * Radical left				−.25
				[−.83; .39]
AIC	9675.21	9676.55	9665.27	9678.42
BIC	9748.11	9760.67	9754.99	9768.15
Num. obs.	2014	2014	2014	2014
N. parties	123	123	123	123
N. countries	25	25	25	25

* 0 outside the bootstrapped confidence interval

Table A.6: Individual and Party-Level Determinants of Percentage Changes in Followers – July 09-13; using tweets from June 09 to Aug. 14

	Model 1	Model 2	Model 3	Model 4
Intercept	5.61*	5.86*	5.92*	5.86*
	[4.40; 6.78]	[4.43; 7.27]	[4.48; 7.28]	[4.51; 7.32]
Male	−.00	−.00	.01	−.00
	[−.24; .23]	[−.24; .24]	[−.25; .26]	[−.26; .26]
Terms in office	−.03	−.03	−.03	−.03
	[−.13; .08]	[−.13; .07]	[−.13; .07]	[−.13; .07]
Cabinet experience	.17*	.17*	.17*	.17*
	[.05; .29]	[.04; .31]	[.04; .31]	[.04; .29]
Twitter Sentiment	−.21	−.22	−.21	−.22
	[−.71; .28]	[−.73; .29]	[−.64; .26]	[−.72; .23]
Twitter EU Sentiment	.05	.05	.01	.06
	[−.10; .20]	[−.10; .19]	[−.14; .16]	[−.09; .21]
Nr. of Followers (log)	−.57*	−.57*	−.57*	−.57*
	[−.66; −.47]	[−.66; −.47]	[−.66; −.47]	[−.66; −.47]
Seat share	.00	.00	.00	.00
	[−.01; .02]	[−.01; .02]	[−.01; .02]	[−.01; .02]
In government	−.19	−.22	−.20	−.22
	[−.70; .28]	[−.72; .28]	[−.67; .27]	[−.73; .29]
EU Position	−.05	−.09	−.10	−.09
	[−.17; .08]	[−.28; .08]	[−.26; .09]	[−.29; .09]
Radical right		−.42	−.52	−.42
		[−1.49; .69]	[−1.49; .57]	[−1.56; .65]
Radical left		−.01	−.02	−.01
		[−.80; .79]	[−.78; .78]	[−.84; .79]
Twitter EU Sentiment * Radical right			.96*	
			[.18; 1.73]	
Twitter EU Sentiment * Radical left				−.23
				[−.84; .32]
AIC	10102.68	10105.43	10101.05	10107.38
BIC	10176.01	10190.04	10191.31	10197.64
Num. obs.	2082	2082	2082	2082
N. parties	127	127	127	127
N. countries	26	26	26	26

* 0 outside the confidence interval

5.2 Dropping outliers

The following table contains the same models from Table 1 in the paper, but dropping observations whose percentage change in number of followers during the purge was larger than 30%.

Table A.7: Individual and Party-Level Determinants of Percentage Changes in Followers Dropping outliers – July 09-13

	Model 1	Model 2	Model 3	Model 4
Intercept	3.36*	4.05*	4.02*	4.05*
	[2.38; 4.28]	[2.81; 5.20]	[2.79; 5.16]	[2.82; 5.16]
Male	−.10	−.10	−.09	−.10
	[−.26; .08]	[−.27; .07]	[−.24; .09]	[−.26; .07]
Terms in office	−.03	−.03	−.03	−.03
	[−.10; .03]	[−.10; .04]	[−.10; .03]	[−.10; .04]
Cabinet experience	.11*	.11*	.11*	.10*
	[.03; .19]	[.02; .19]	[.02; .19]	[.02; .19]
Twitter Sentiment	−.61*	−.62*	−.64*	−.63*
	[−.98; −.24]	[−.96; −.26]	[−1.00; −.27]	[−1.01; −.26]
Twitter EU Sentiment	.03	.03	−.04	.04
	[−.07; .13]	[−.08; .12]	[−.14; .07]	[−.07; .15]
Nr. of Followers (log)	−.35*	−.35*	−.35*	−.35*
	[−.41; −.28]	[−.41; −.29]	[−.41; −.29]	[−.41; −.29]
Seat share	.00	.00	.00	.00
	[−.02; .02]	[−.01; .02]	[−.01; .02]	[−.01; .02]
In government	−.26	−.32	−.23	−.32
	[−.68; .18]	[−.76; .12]	[−.67; .18]	[−.76; .10]
EU Position	.06	−.06	−.06	−.06
	[−.06; .17]	[−.23; .10]	[−.21; .10]	[−.22; .10]
Radical right		−1.04*	−1.16*	−1.04*
		[−2.05; −.11]	[−2.06; −.24]	[−1.97; −.10]
Radical left		−.18	−.13	−.17
		[−.89; .56]	[−.80; .56]	[−.89; .51]
Twitter EU Sentiment * Radical right			1.86*	
			[1.30; 2.42]	
Twitter EU Sentiment * Radical left				−.20
				[−.64; .20]
AIC	7665.12	7664.35	7623.57	7666.76
BIC	7737.35	7747.70	7712.48	7755.67
Num. obs.	1913	1913	1913	1913
N. parties	119	119	119	119
N. countries	25	25	25	25

* 0 outside the confidence interval

6 Difference-in-Differences Estimates

As a robustness test, we can estimate whether certain party families lost more followers due to the purge with a Difference-in-Differences (DiD) model. The unit of observation is the tweet, and the dependent variable is the absolute number of followers of that user at the moment of posting it. We estimate the following model for each party family:

$$Followers_{ijk} = \beta 0_{00} + \beta 1 * Time_{ijk} + \beta 2 * Family_{ijk} + \beta 3 * Time_{ijk} * Family_{ijk} + \epsilon_{ijk} + \upsilon_{jk} + \nu_k \quad (1)$$

Where $Followers_{ijk}$ is the number of followers, varying at the tweet (i), user (j), and country (k) levels. $\beta 0$ is the grand mean of followers, $\beta 1$ the main effect of the $Time$ dummy, meaning whether a tweet was posted before (0) or after (1) the purge; $\beta 2$ the main effect of the party $Family$ dummy (we run the model seven times, one time with each party family determined as 1 and the others as 0), and $\beta 3$ the DiD estimate of the effect of party family before/after the purge. ϵ_{ijk} is the between-tweets residual variance, υ_{jk} is the between-users residual variance, and ν_k is the between-country residual variance.

Table A.8: Difference-in-Differences Estimates on Change in number of Followers – July 09-13; Using Tweets from July 01 to July 21

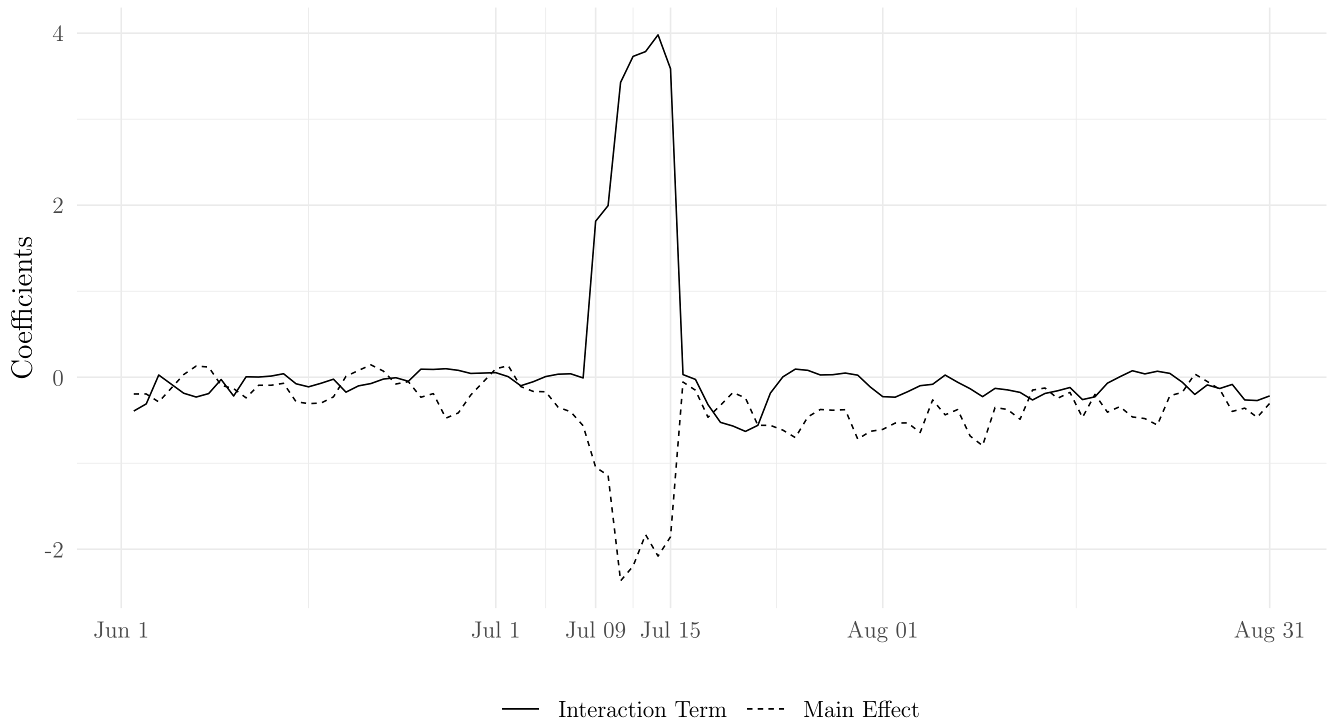
	Rad. Right	Rad. Left	Christ. Dem.	Conservative	Socialist	Regionalist	Green
(Intercept)	18563.68^{***}	17760.44^{***}	20597.62^{***}	21037.62^{***}	18854.11^{***}	20279.32^{***}	20066.62^{***}
	(3217.22)	(3117.51)	(3081.26)	(3288.60)	(3245.09)	(3147.62)	(3105.31)
time	-109.21^{***}	-229.19^{***}	-238.70^{***}	-269.42^{***}	-255.94^{***}	-244.37^{***}	-233.90^{***}
	(21.89)	(22.74)	(22.21)	(24.85)	(25.43)	(22.36)	(22.16)
Family	21282.00^{*}	25578.45^{***}	-7785.45	-5960.75	4556.63	-9863.13	-2801.88
	(9624.02)	(7569.79)	(8511.75)	(4958.44)	(4625.04)	(9287.23)	(9831.45)
time:Family	-3251.52^{***}	13.91	242.78^{*}	174.96^{***}	103.14^{*}	286.67^{**}	146.87
	(114.53)	(76.57)	(105.60)	(51.05)	(48.83)	(93.44)	(110.26)
Num. obs.	40704	40704	40704	40704	40704	40704	40704

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

7 Placebo Tests

To make sure we are not capturing a momentary fluke, we have used a placebo test approach. We attributed the purge to have happened in each day between June 01 and August 31 – thus calculating the percentage changes in followers with the same windows of days before and after each one, and fit models 1 and 3 each time to each of the new calculation of followers ratios. The dashed line in Figure A.3 shows the main coefficient of radical right from Model 1 for each day. We see that the significant negative result only happens around the time of the actual purge: between July 09 and 15. If we were to pretend it happened any day before or after that, never are radical right parties associated with such large drops in the numbers of friends and followers. The solid line shows the interaction effect between radical right and EU sentiment: once again, the large effects on the number of followers happen only around the purge, never before or after.

Figure A.3: Placebo tests: Purge happening each day between June 01 and August 31



8 Effects on Friends Counts

The following models are the same as those in Table 1 of the main paper, but using percentage changes in the friends counts (who an account follows) between July 09 and 14, instead of percentage changes in follower counts.

Table A.9: Individual and Party-Level Determinants of Percentage Changes in Friends – July 09-13

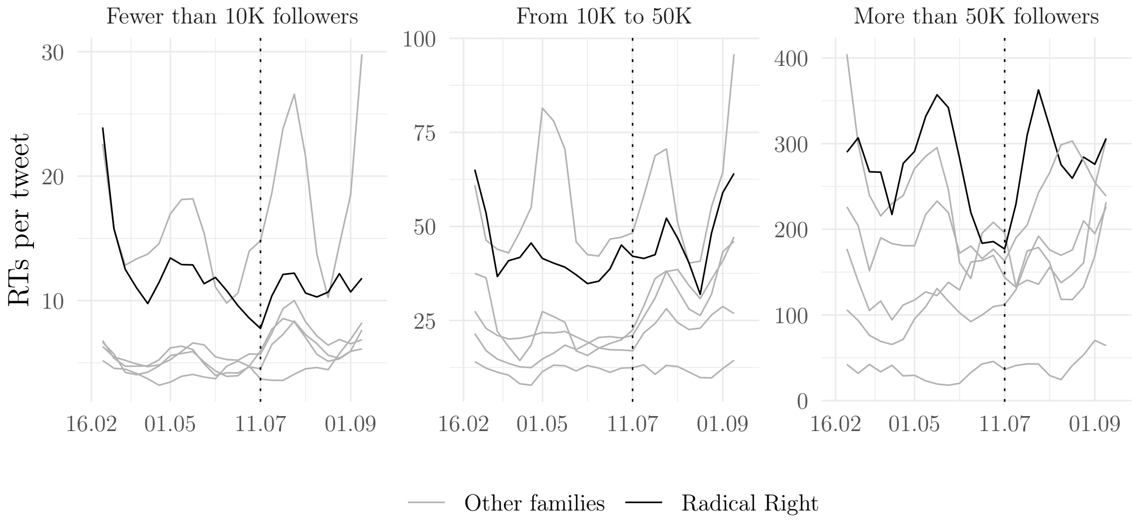
	Model 1	Model 2	Model 3	Model 4
Intercept	−.09	.45	.45	.45
	[−1.19; 1.00]	[−.88; 1.72]	[−.84; 1.83]	[−.80; 1.80]
Male	.11	.11	.11	.11
	[−.14; .35]	[−.14; .35]	[−.14; .36]	[−.13; .38]
Terms in office	−.00	−.00	−.00	−.00
	[−.09; .10]	[−.10; .09]	[−.10; .09]	[−.10; .09]
Cabinet experience	.02	.02	.02	.02
	[−.11; .16]	[−.12; .14]	[−.11; .14]	[−.10; .14]
Twitter Sentiment	.01	−.01	−.01	−.01
	[−.42; .47]	[−.52; .47]	[−.53; .48]	[−.49; .47]
Twitter EU Sentiment	−.04	−.04	−.02	−.03
	[−.19; .12]	[−.19; .13]	[−.19; .15]	[−.20; .13]
Nr. of Followers (log)	−.08	−.09	−.09	−.09
	[−.18; .01]	[−.18; .01]	[−.19; .01]	[−.19; .01]
Seat share	.01	.01	.01	.01
	[−.01; .02]	[−.01; .02]	[−.01; .02]	[−.01; .02]
In government	.11	.10	.08	.09
	[−.39; .60]	[−.44; .55]	[−.43; .55]	[−.41; .62]
EU Position	.13*	.04	.04	.04
	[.00; .25]	[−.12; .20]	[−.11; .20]	[−.12; .21]
Radical right		−1.04	−1.00	−1.04*
		[−2.08; .04]	[−2.03; .10]	[−2.06; −.03]
Radical left		.21	.21	.21
		[−.53; .98]	[−.49; 1.02]	[−.59; .98]
Twitter EU Sentiment * Radical right			−.52	
			[−1.35; .31]	
Twitter EU Sentiment * Radical left				−.11
				[−.77; .52]
AIC	9220.71	9218.74	9219.14	9221.05
BIC	9292.95	9302.10	9308.05	9309.96
Num. obs.	1914	1914	1914	1914
N. parties	119	119	119	119
N. countries	25	25	25	25

* 0 outside the confidence interval

9 Number of Retweets

Figure A.4 is created based on Retweet counts collected in October 2018 for all tweets posted by MPs which were not replies or retweets themselves. These are weekly moving averages (3-weeks window) to smooth some of the random variance.

Figure A.4: Weekly Moving Averages of Retweets per Post



10 Bot Detection Algorithm

We used the R package `tweetbotornot` (Kearney, 2020) on a random list of 20,799 followers of French politician Marine Le Pen and 19,888 followers of German politician Christian Lindner, to get an estimate of how many bots might follow these politicians using a different method than the Twitter purge. These lists were based on their total follower lists in early April 2020. The algorithm uses only users' information and activity patterns to predict the probability of an account being a bot, and is a machine learning model trained on various sources of labelled data. These two politicians were chosen to illustrate the point because they are highly popular, both on the right side of the ideological spectrum, and both in opposition to a center-right national government. However, Le Pen belongs to a radical right party, while Lindner belongs to a center-right liberal party.

We find that the algorithm is likely to overestimate the proportion of bots: 67.5% of the ca. 20,000 random Le Pen's followers had a probability of being bots above 0.5, against 54.98% of Lindner's followers. Regardless, we still observe a higher probability of bot followers for the radical right politician than the center right.

References

Kearney, Michael W. 2020. *tweetbotornot: Classify Twitter Users as Bot or Not.* R package version 0.1.0.
URL: *https://tweetbotornot.mikewk.com*