Political polarization dynamic during the covid-19 in developed countries

Juan Diego Luksic*, Bruno Chaves Ferreira, Eric Brousseau* ****

*Governance and Regulation Chair, Paris Dauphine—PSL University
*Applied Computational Social Science Institute (ACSS), PSL University, Paris

January 31, 2024

(Work in Progess)

Abstract

This paper studies how the government-opposition political polarization changed after nine developed countries decided to implement lockdowns during the onset of the covid-19 pandemic. Using Twitter data from party, party political leaders, and government official accounts, we estimate the political polarization evolution during 2020 and provide validation evidence by observing recent presidential elections in the US and Poland and studying the effect of the protests after George Floyd's murder. We also show that the high level of political polarization is not driven by different communication styles between government officials and party political leaders. Implementing a stacked event study, we show that politicians' tweet is 2.5 percentage points less likely to identify group belonging two weeks after a lockdown is implemented. This result vanishes four weeks after the lockdown. Importantly, results are similar when focusing on Covid-19-related tweets.

1 Introduction

The covid-19 posed unprecedented challenges to political leaders across the world. Most public health experts agree on the importance of restrictive measures such as lockdowns to slow the spread of the virus, ease the burden on the healthcare system and save lives. The political environment plays a key role in the decision-making process affecting its efficiency and efficacy. Political polarization can delay fiscal stabilization, increase uncompromising obstructionism, political gridlock, the politicization of regulators institutions and policy uncertainty with implications for government functioning and macroeconomic performance (Baker et al., 2014; Davis, 2019; Mian et al., 2014; Pastor and Veronesi, 2012). Also, there is literature providing evidence on how the financial crisis and elections can increase political polarization (Baker et al., 2020; Mian et al., 2014). Eventually, the political polarization can also be affected by the pandemic crisis or by the implementation of life-disturbing policies such as a domestic lockdown. Therefore, two questions arise in the interaction between political polarization and the policy decision process during the pandemic. First, how the political polarization affects the government's decision-making process, and second, how stringent policies affect political polarization.

This paper aims to answer the second question by evaluating how lockdown decision affects political polarization at the onset of the pandemic and by constructing and validating a political polarization indicator using Twitter political accounts. Using tweets of political parties, party leaders and government officials from 12 developed countries collected by data scientists of the Governance and Regulation Chair, Paris Dauphine—PSL (henceforth the Chair), we estimate political polarization weekly following a leave-out estimate (Gentzkow et al., 2019). In contrast to political polarization using Twitter users, Twitter political accounts have the advantage of better representing the political debate at the country level. Twitter users are not representative of the population because they can choose to be on the platform or not. In the case of political actors, it is not compulsory, but they cannot waste the opportunity to use the platform to communicate their ideas and their work.[1]

Political polarization estimates using Twitter political accounts are higher than using congress speeches in the US (ibid.). This is also true when comparing political polarization using the Twitter community for specific events (Demszky et al., 2019). We show that this finding is not driven by a different communication style between public officials and political party leaders. We then validate the indicator by focusing on the US and Poland presidential elections in 2020. Then we validate the leave-out estimates by showing that the protests after George Floyd's murder had a positive effect on the polarization estimate. Finally, using a stacked event study, we found that the lockdown reduced political polarization, but this effect vanished after five weeks. The next section describes the data used for the analysis, section 4 describes the empirical analysis and results, and section 5 concludes.

2 Data

2.1 Political Twitter accounts

The polarization analysis focuses on western developed countries. We selected 12 countries: Poland, the United States, Germany, France, Sweden, Spain, the United Kingdom, Italy, Canada, Netherlands, Australia, and New Zealand. We use the dataset named Parlgov (Döring and Manow, 2022) to extract all political parties in the last lower house elections with more than 1% of vote share. Then we selected the political leaders of each party[2] and the ministries of the government. One caveat of Parlgov is that it does not include the US. Therefore, in this case, we selected all accounts by hand following the next criteria for Republicans and Democrats: chairperson of the party, the speaker, the majority leader and the minority leader of the house of representatives, and presidential candidates (2020 elections). Data scientists from the Chair scrapped tweets from 1st December 2019 to 28th January 2021. Importantly, because most tweets were written in local languages, they were translated to English with Google Translation API.

After scrapping the tweets from these accounts, we end up with a total of 389,297 tweets. The dataset contains 357 Twitter accounts: 100 official accounts of parties, 82 accounts of party leaders and 175 accounts of

[1] In the collected data 82% of party leaders have active Twitter accounts during the period.

[2] Canen et al., 2021 show the importance of party leaders to drive political polarization.

ministers and ministries. In this study, political polarization is estimated using two groups: government and opposition. The government group is not only the members of the government but also the parties of the government coalition. The opposition comprises all the parties that do not belong to the government coalition. In total, we have 234 Twitter government accounts and 123 Twitter opposition accounts. Table 1 shows the distribution of tweets by country and governmentopposition groups. In most countries, the proportion is equivalent, though there are some exceptions, like Germany and Italy. Also, not all countries have the same quantity of tweets leading to a different level of reliability of the leave-out estimate.

2.2 Public measurements and COVID-19 dynamic

We took the public measurements from Response2covid19 dataset (Porcher, 2020). This data collects information from several datasets and report daily public measures in response to covid-19 in more than 200 countries identifying public health policies (domestic lockdown, school closure, curfew, postponed elections, among others) and economic interventions (tax credits, tax delays, wage support among others).

The covid-19 dynamic across countries is collected from COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (Dong et al., 2020). This data contains daily information on infection and deaths at the country level.

3 Quantifying political polarization

3.1 Methods

The estimation of the political polarization dynamics within countries follows the method from Gentzkow et al., 2019. The advantage of this indicator is its relative simplicity to implement, the possibility to estimate standard errors and implement placebo tests, and that belongs to the supervised estimators, which in the case of political polarization seems to work better (Goet, 2019). To estimate this index with Twitter data, we clean tweets and then proceed with estimation.

3.1.1 Data cleaning

We proceed by creating a vocabulary for each period. We select unigrams and bigrams and remove links, emails, digits, hashtags, punctuation and stopwords. Then implement stemming via SnowballStemmer. Also, we keep words occurring more than ten times in a period/country and more than 200 times in the data (pooling countries and periods). We group tweets by account and day and define the period of analysis in one week.

3.1.2 Leave-out political polarization estimation

The polarization index proposed in Gentzkow et al., 2019 measure the expected posterior probability that a speech is from the true party after a single token drawn at random from the tweets produced by the tweeters. The estimation is as follows:

$$\Pi^{LO} = \frac{1}{2}\left(\frac{1}{|D|}\sum_{i\in D}\hat{q}_i\hat{\rho}_{-i} + \frac{1}{|R|}\sum_{i\in R}\hat{q}_i(1-\hat{\rho}_{-i})\right) \tag{1}$$

where $\hat{q}_i = c_i/m_i$ is the vector of empirical token frequencies for tweeter i, with c_i being the vector of token counts for tweeter i and m_i the sum of token counts for tweeter i; and $\hat{\rho}_i = (\hat{q}_i^{D\setminus i} \oslash (\hat{q}_i^{D\setminus i} + \hat{q}_i^{R\setminus i}))$ is a vector of empirical posterior probabilities, excluding speaker i and any token that is not used by at least two speakers. $\oslash$ is a element-wise division and $\hat{q}^G = \sum_{i \in G} c_i / \sum_{i \in G} m_i$. Note that the leave-out estimator (ρ) helps to address the finite sample bias due to the mechanical correlation of the sampling errors of q and ρ.

3.2 Leave-out estimates

We estimate the political polarization index in 9 countries weekly. We discarded Australia, New Zealand and Italy because there were too few observations in some periods. Figure 1 shows the dynamics of the index in weeks from December 2019. The black line shows the polarization index using government opposition. The blue line shows a placebo to quantify the noise by assigning groups randomly. The red dotted line is fixed in 0.5. The value resulting from the placebo random group assignment (green lines) are all close to 0.5, suggesting that the actual values are not a result of noise. Results show that countries display different levels of polarization; for example, Poland and the United States show higher levels (mean *approx* 0.6 over the period observed). This could be a cultural difference or the fact that these countries face higher levels of polarization in this period of analysis. In the case of the United States, recent studies have shown that the level of polarization has been higher in recent years (Gentzkow et al., 2019). Importantly, the levels of the polarization index using Twitter political data are higher than analyzing US congress speeches or Twitter community analyzing mass shooting eventsDemszky et al., 2019. In our case, the average is ~ 0.57, while in the case of Gentzkow et al., 2019, recent political polarization is ~ 0.53. It is not clear why the index is significantly higher in the Twitter political context. One reason could be that the tweet limit of characters - up to 280 - pushes politicians to choose more appealing words for their political sector, resulting in a text containing words that are more "purely" in their political stands. Another reason could be that the difference is because we am comparing government with opposition and, therefore, mostly ministers with parties and party leaders. To address this alternative, we evaluate if the results are similar when using left-right using only party and party leaders[3]. The grey line of Figure 1 shows this exercise, and graphs depict a similar level for both indicators of political polarization with the exception of Poland. The case of Germany can be explained because of the great imbalance of the number of accounts in government-opposition shown in table 1, which makes it a less stable indicator when grouping differently. Overall, these results suggest that the higher level of political polarization in government opposition is not explained because public officials use a different communication style.

3.3 Major political events in US and Poland

We validate the political polarization indicator in this setting by mapping the major political events happening from the end of 2019 to the beginning of 2021. We focus on Poland and the US because both had election processes during the period observed,

[3]The left-right classification is done with the Parlgov dataset that assigns a time-invariant unweighted mean value of information from party expert surveys on a 0 to 10 scale in four subcategories.

and therefore identification of major political events is straightforward. We selected different events that are supposed to convey higher political polarization. In the case of the US, relevant election-related events include the impeachment of the president and the murder of George Floyd. Election-related events consider the episodes occurring from the presidential nominating convention (17th of August) to the Capitol attack (6th of January). We also added Christmas, Thanksgiving and the 4th of July as events that should engage sentiments of national unity. Similarly, in Poland, we selected relevant election-related events and the Polish Constitutional Tribunal's decision to make abortion unconstitutional. Election-related events consider episodes occurring from the parliament announcement of the presidential elections (5th of February) to the challenge of the results from the opposition (16th of July). We also added Christmas and Independence Day as events that should engage unity and decrease polarization. Figure 5 shows the polarization dynamics in the US. From December 2019 to February 2020, the political polarization is high (≈ 0.6), probably due to the impeachment in the House of Representatives and in the Senate, with the exception of Christmas and the week after the impeachment in the Senate. Then polarization stayed low until the week after the murder of George Floyd, probably as a consequence of the massive protests all over the country organized by "Black Lives Matter." The indicator spikes again with the presidential nominating convention and stays at high levels until the elections. Then, polarization goes up and down, probably due to high polarization in protest of electoral fraud and the Capitol attack and low polarization in Thanksgiving and Christmas. Figure 3 shows the polarization dynamic in Poland. Political polarization remains at a high level (≈ 0.6) -except for Christmas- and decays slowly until April. Polarization started to increase because the government persevered in maintaining the presidential elections, even though NGOs and the opposition were in favor of postponing the elections. The government proposed postal voting but ended up failing and postponed the election at the peak of political polarization. Then in June, political polarization decreased until the former president was reelected at the beginning of July, and the opposition challenged the election. After this, we see ups and downs in polarization until the WSA court (administrative court) determines that the postal voting bill violated the law, and the opposition demands the prime minister's resignation. In October, we observed another peak related to the decision to make abortion unconstitutional. Finally, we observe a decay in the polarization for the rest of the year, with clear plunges during independence day and Christmas week.

3.4 Protests

The previous analysis is illustrative of what the indicator is capturing, but to have a systematic approximation, we observe the protests occurring in the developed countries after the murder of George Floyd and test if they match with higher levels of leave-out political polarization estimates. We choose the murder of George Floyd because it is unexpected to politicians, and the event is unrelated to the levels of political polarization of each country. To observe the protests, we use Carnegie's global protest tracker (Tracker, 2020) and select the countries in the sample. In the period, four protests occurred in the nine developed countries (USA, France, United Kingdom and Germany). Figure 4 shows the coefficients of a stacked event study[4] showing that the impact of the protests on political polarization is positive but starts the week after the protests and remains high for four weeks. The increase in the week after the protest is 2.5 percentage points (hereafter pp.), i.e.; it is 2.5pp (statistically significant at 10% level, wild bootstrap Cameron et al., 2008) more likely to identify a politician speech based on their group belonging (majority-opposition). On average, the increase four weeks after the protests is 1pp (statistically significant at 10% level, wild bootstrap ibid.).

4 Empirical approach and results

4.1 Empirical approach

We evaluate the impact of lockdown on political polarization using the first lockdown in the first semester of 2020. We focus on this policy because it is a high-stakes policy and on this period because there was no precedence of lockdown due to the pandemic, thus, more uncertainty. Most countries in the sample implemented a lockdown in March 2020, but some did not (Canada, Netherlands and Sweden). Using this variation, we evaluate if implementing a lockdown decreases polarization. De Chaisemartin and d'Haultfoeuille, 2020 show that two-way fixed effects estimates are biased under heterogeneity treatment effect in a staggered design because it suffers from negative weights. To address this problem, we implement a stacked event study approach (Cengiz et al., 2019; Deshpande and Li, 2019). This approach consists of implementing separate experiments using "clean" controls, always untreated units, and units that implemented lockdown in the same period. Then, the experiments are stacked using the following specification:

$$y_{itd} = \alpha_i d + \theta_t d + \beta_{\tau_{id} t} + \gamma X_{itd} + \epsilon_{itd} \quad (2)$$

where y_{itd} is the leave-out estimate in country i in time to the event t in experiment d, $\alpha_i d$ is a country i experiment d fixed-effect and $\theta_t d$ is time to event t experiment d fixed-effect, $\beta_{\tau_{id} t}$ is a coefficient varying with time to event t and and treatment τ_{id} a variable varying with country i and experiment d. Finally, X_{itd} are covariates varying in time to event t, country i and experiment d, and ϵ_{itd} is the error term that correlates at country level.

This strategy requires the parallel trend assumption, i.e., in the absence of treatment, control, and treated countries would follow a similar trend. The lockdown implementation is endogenous, but we can control after variables that possibly drive the decision, such as deaths and demographic distribution. In other words, our strategy assumes that unobservables driven differential trends are orthogonal to the treatment.

Finally, to account for precision in the leave-out estimates, we take the inverse of the placebo standard deviation and run the regression weighting by these levels. Also, because there are only nine countries, clustered standard errors at the country level are obtained after wild bootstrap (Cameron et al., 2008).

[4] I described the methodology in Section 3.1. Also, it is stacked because the protest in France and Germany were one week after the protest in USA and United Kingdom.

4.2 Results

Figure 5 depicts the effect of lockdown on the leave-out political polarization estimated weekly. Before $t - 1$ treated and control countries follow a similar trend in political polarization, suggesting evidence in favor of the parallel assumption. We observe a decrease that reaches the lowest level of 3.5pp in the third week of implementation (significant at 10% level using wild cluster bootstrap at country level). This suggests that after the lockdown implementation, politicians use more similar words, and therefore the probability of predicting group-belonging base on speech reduces. Then the effect vanishes toward the fifth week of implementation.

4.3 Differences in topic

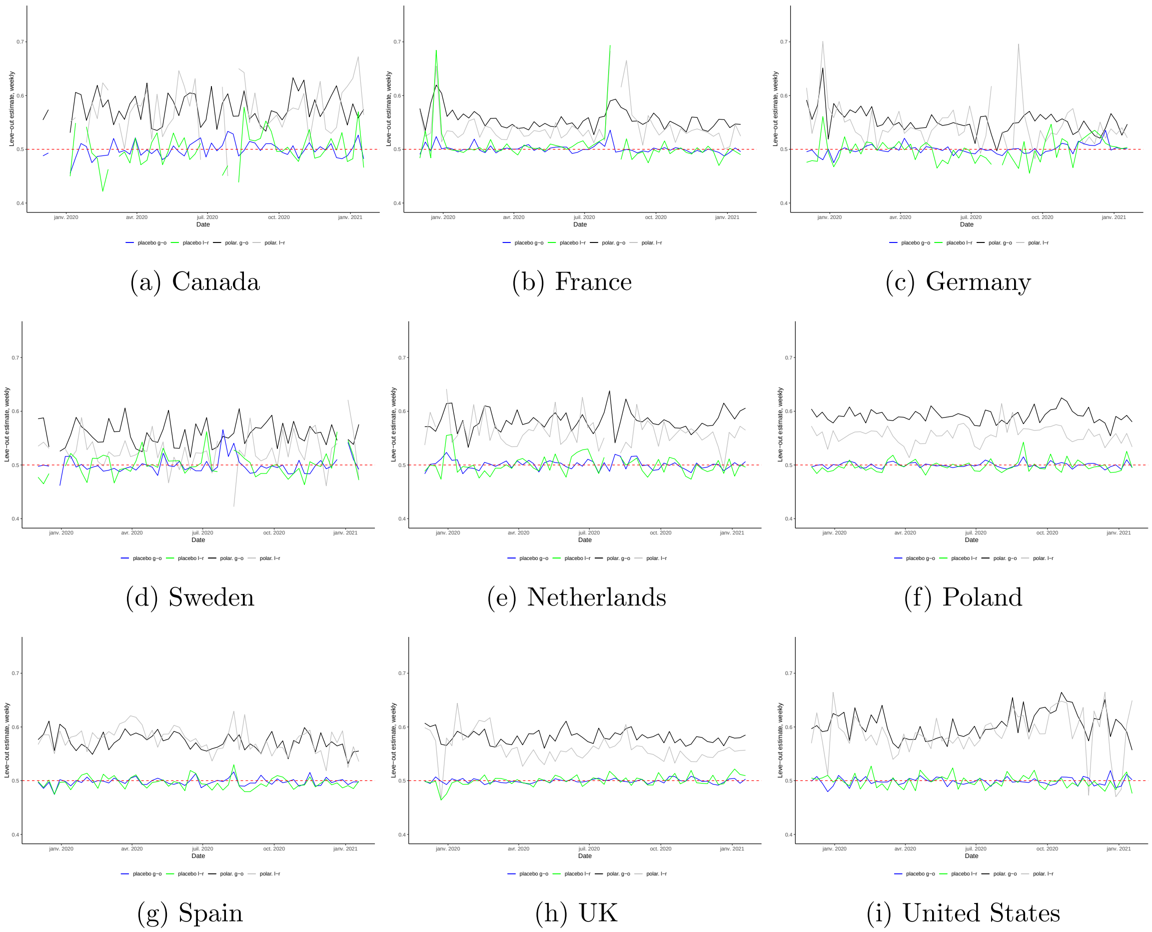
The leave-out indicator exploits the inter and intra-group word variation in the tweets. Therefore, a higher level of polarization could be related to different rhetoric when discussing the same or different topics. To address this question, we evaluate the same regression by choosing tweets related to the pandemic using the Covid19 dictionary in Table 2 in the appendix. Figure 6 shows that results are very similar, suggesting that discussions on a different topic did not drive the decrease in political polarization.

5 Discussion

We studied the political polarization dynamic of nine developed countries during the onset of the pandemic. We first validate the leave-out estimate (Gentzkow et al., 2019) using Twitter political accounts by studying the US and Poland elections and observing the protest dynamic after George Floyd's murder. Then we evaluate the effect of lockdown on political polarization, finding that there is a negative effect that vanishes after a month. This suggests that stringent policies, such as lockdowns in the context of the pandemic, can bring short-term moments of low political polarization, similar to dates that convey unity, such as National Day or Christmas. Interestingly, it would be relevant to evaluate if the nature of the policy drives this effect - politicians' lockdown forces them to stand by their political priorities- or the positive effect of the policy on reducing infection. In any case, these results support public health stringent policies adoption when needed by suggesting no detrimental effect in the political environment.

Figures

Figure 1: Polarization index dynamic in nine developed countries



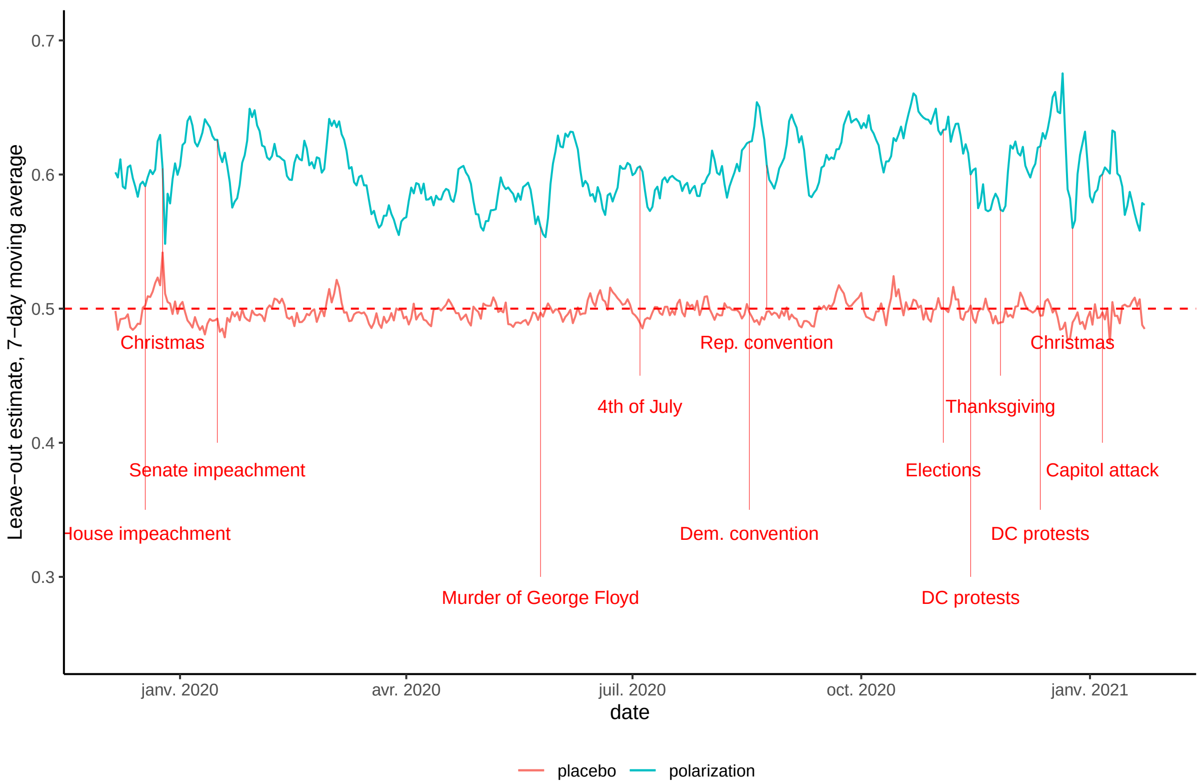
(a) Canada (b) France (c) Germany

(d) Sweden (e) Netherlands (f) Poland

(g) Spain (h) UK (i) United States

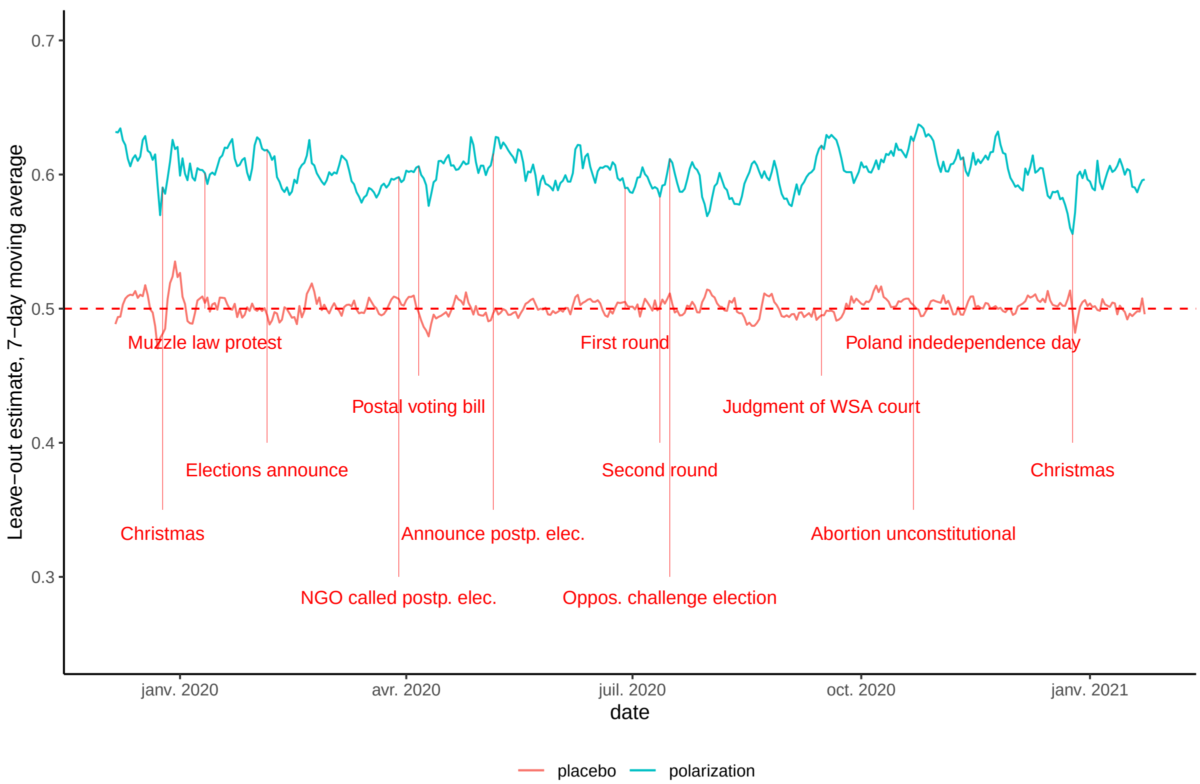
Note: Leave-out estimate for developed countries estimated weekly following the method described in Section 3.1. The black line depicts the leave-out estimate using majority-opposition and the blue line its placebo . The grey line depicts the leave-out estimate using left-right and the green line its placebo.

Figure 2: United States political polarization dynamic and major political events



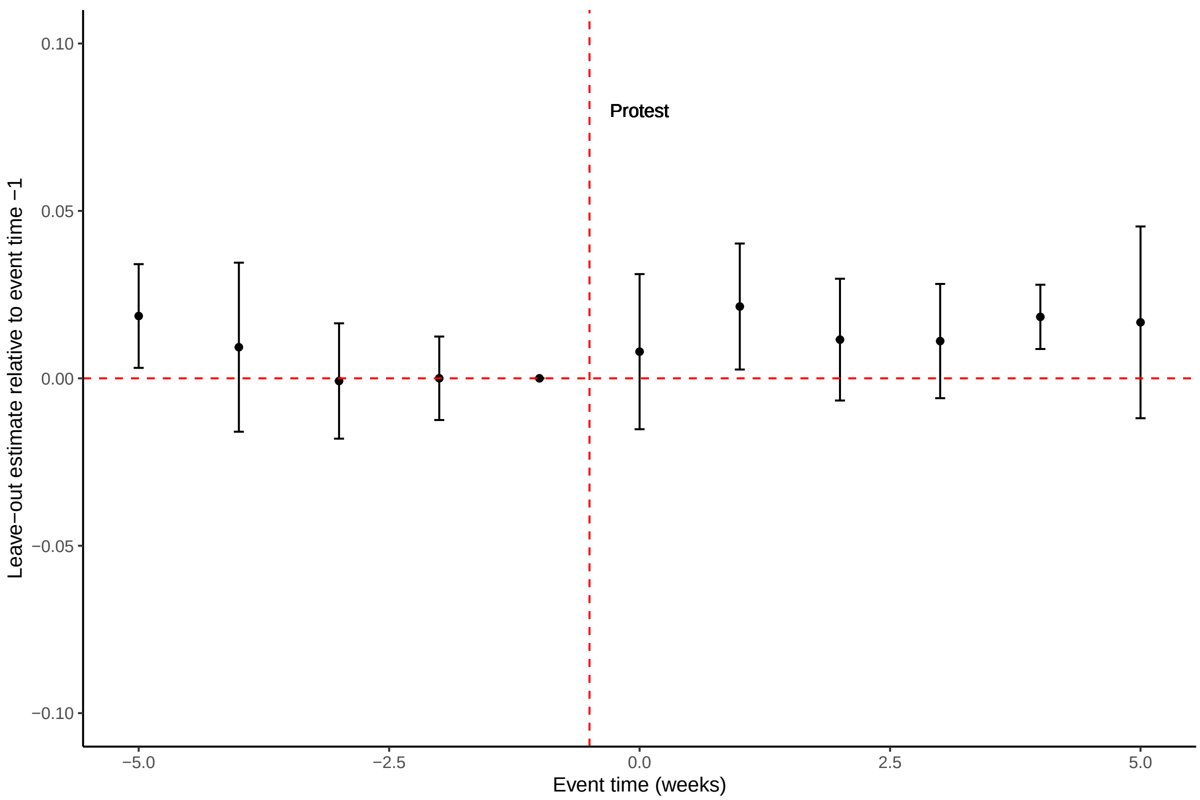
Note: Leave-out estimate for the US estimated in 7-day moving average following the method described in Section 3.1.

Figure 3: Poland political polarization dynamic and major political events



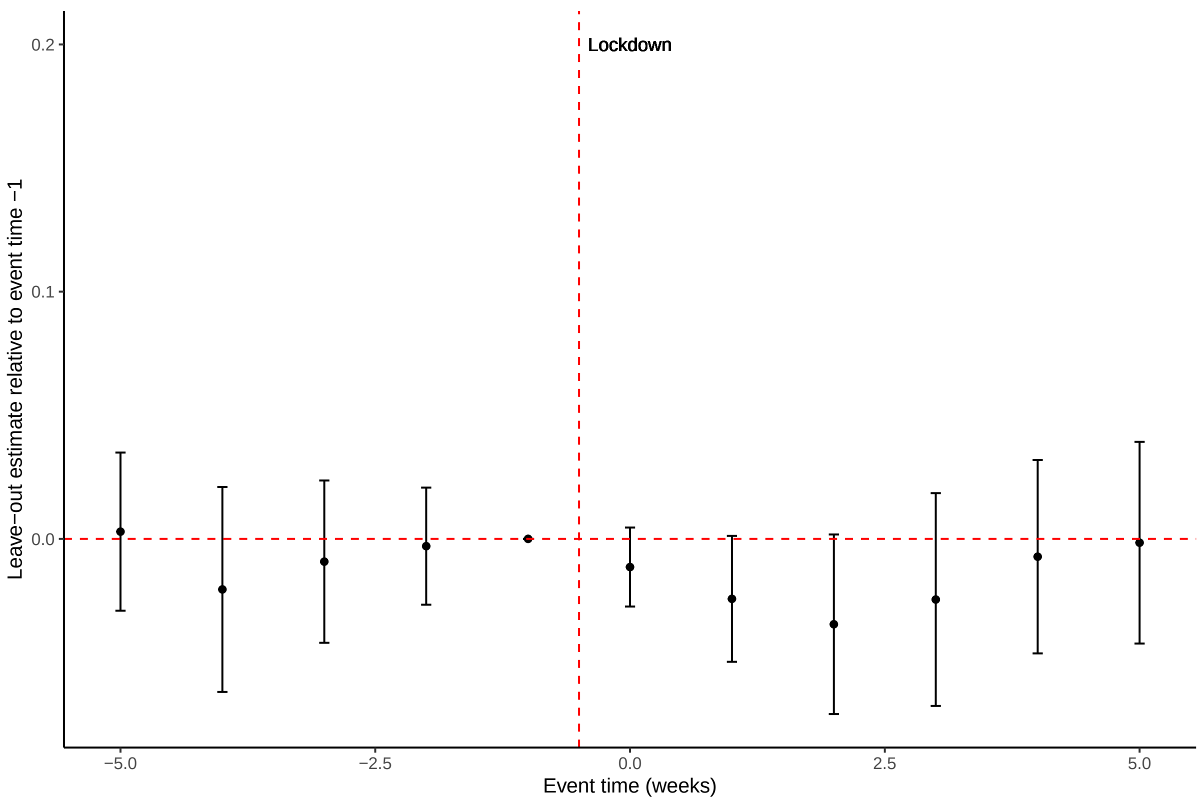
Note: Leave-out estimate for Poland estimated in 7-day moving average following the method described in Section 3.1.

Figure 4: Stacked event study for protests after George Floyd murder



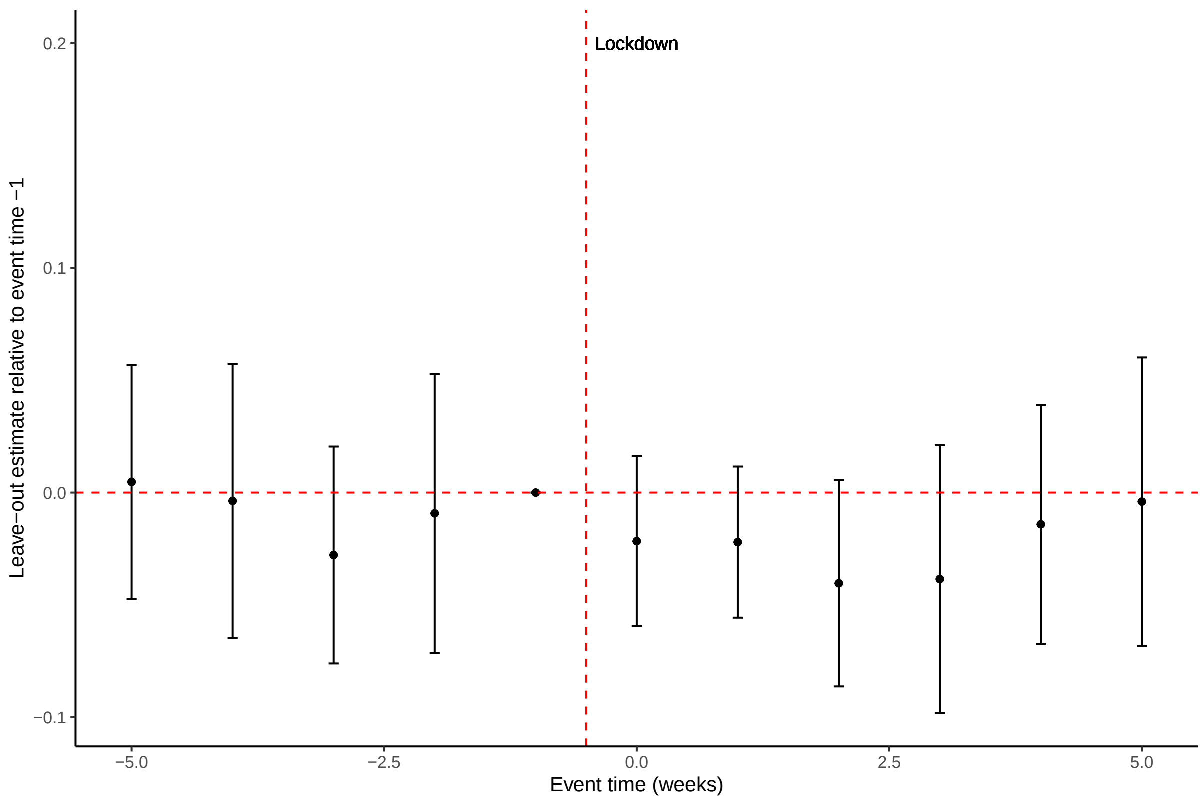
Note: Leave-out estimate estimated weekly following the method described in Section 3.1.

Figure 5: Stacked event study for lockdown



Note: Leave-out estimate estimated weekly following the method described in Section 3.1.

Figure 6: Stacked event study for lockdown considering Covid-19 related topics



Note: Leave-out estimate estimated weekly following the method described in Section 3.1.

Tables

Table 1: Descriptive statistics of number of tweets by country and group

Country		Opposition	Government	All
Australia	N	3105	4753	7858
	% row	39.5	60.5	100.0
Canada	N	9608	23857	33465
	% row	28.7	71.3	100.0
France	N	16931	21584	38515
	% row	44.0	56.0	100.0
Germany	N	5397	29604	35001
	% row	15.4	84.6	100.0
Italy	N	1514	8604	10118
	% row	15.0	85.0	100.0
Netherlands	N	10105	16331	26436
	% row	38.2	61.8	100.0
New Zealand	N	2030	4841	6871
	% row	29.5	70.5	100.0
Poland	N	41902	31441	73343
	% row	57.1	42.9	100.0
Spain	N	18094	26848	44942
	% row	40.3	59.7	100.0
Sweden	N	7419	5430	12849
	% row	57.7	42.3	100.0
United Kingdom	N	45586	24350	69936
	% row	65.2	34.8	100.0
USA	N	8861	21102	29963
	% row	29.6	70.4	100.0
All	N	170552	218745	389297
	% row	43.8	56.2	100.0

Note: Number of tweets by country between December 2019 and January 2021 for the Twitter accounts selected.

References

Baker, Scott R, Aniket Baksy, Nicholas Bloom, Steven J Davis, and Jonathan A Rodden (2020). *Elections, political polarization, and economic uncertainty.* Tech. rep. National Bureau of Economic Research.

Baker, Scott R, Nicholas Bloom, Brandice Canes-Wrone, Steven J Davis, and Jonathan Rodden (2014). "Why has US policy uncertainty risen since 1960?" In: *American Economic Review* 104.5, pp. 56–60.

Cameron, A Colin, Jonah B Gelbach, and Douglas L Miller (2008). "Bootstrap-based improvements for inference with clustered errors". In: *The review of economics and statistics* 90.3, pp. 414–427.

Canen, Nathan J, Chad Kendall, and Francesco Trebbi (2021). *Political Parties as Drivers of US Polarization: 1927-2018.* Tech. rep. National Bureau of Economic Research.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer (2019). "The effect of minimum wages on low-wage jobs". In: *The Quarterly Journal of Economics* 134.3, pp. 1405–1454.

Davis, Steven J (2019). *Rising policy uncertainty.* Tech. rep. National Bureau of Economic Research.

De Chaisemartin, Clément and Xavier d'Haultfoeuille (2020). "Two-way fixed effects estimators with heterogeneous treatment effects". In: *American Economic Review* 110.9, pp. 2964–96.

Demszky, Dorottya, Nikhil Garg, Rob Voigt, James Zou, Matthew Gentzkow, Jesse Shapiro, and Dan Jurafsky (2019). "Analyzing polarization in social media: Method and application to tweets on 21 mass shootings". In: *arXiv preprint arXiv:1904.01596.*

Deshpande, Manasi and Yue Li (2019). "Who is screened out? Application costs and the targeting of disability programs". In: *American Economic Journal: Economic Policy* 11.4, pp. 213–48.

Dong, Ensheng, Hongru Du, and Lauren Gardner (2020). "An interactive web-based dashboard to track COVID-19 in real time". In: *The Lancet infectious diseases* 20.5, pp. 533–534.

Döring, Holger and Philip Manow (2022). "Parliaments and governments database (ParlGov): Information on parties, elections and cabinets in modern democracies". In: *Development version.*

Gentzkow, Matthew, Jesse M Shapiro, and Matt Taddy (2019). "Measuring group differences in high-dimensional choices: method and application to congressional speech". In: *Econometrica* 87.4, pp. 1307–1340.

Goet, Niels D (2019). "Measuring polarization with text analysis: evidence from the UK House of Commons, 1811–2015". In: *Political Analysis* 27.4, pp. 518–539.

Mian, Atif, Amir Sufi, and Francesco Trebbi (2014). "Resolving debt overhang: Political constraints in the aftermath of financial crises". In: *American Economic Journal: Macroeconomics* 6.2, pp. 1–28.

Pastor, Lubos and Pietro Veronesi (2012). “Uncertainty about government policy and stock prices”. In: *The journal of Finance* 67.4, pp. 1219–1264.

Porcher, Simon (2020). “Response2covid19, a dataset of governments’ responses to COVID-19 all around the world”. In: *Scientific data* 7.1, pp. 1–9.

Tracker, Global Protest (2020). *Global Protest Tracker website.*

A Figures and tables

Table 2: Dictionary of terms Covid19-related

Treatment terms =	Vaccination, Test ,Testing ,Hydroxychloroquine, Reanimation, Spread,
	Vaccine, Vaccines, Spreading, Asymptomatic, Symptomatic, Masks
Emergency measures =	Quarantine, Distancing, Crisis, Emergency,
	lockdown, Lock, confinement, Immunity
Medical institutions =	Hospitals, Hospital, Laboratory, Laboratories, Home