Conflicts Rooted in Socio-Cultural Animosities: A Historical Analysis Based on the SPEED Dataset Covering 1946-2005

Abstract. A vast literature, data projects, and scholarly debates have been coalescing around ethnic conflicts. Yet the notion of *ethnic* remains imprecise, and its operationalizations also endorse various entities of different scopes. The broadest sense *ethnic* includes communal features beyond ethnicity *per se*, such as religious, regional, tribal, racial specifics. The heavily debated core issue of ethnic conflict research is whether ethnic differences lead to conflictual behavior on their own, or only in certain circumstances, such as in the presence of inequalities or of unscrupulous political entrepreneurs exploiting differences. The issue is not esoteric: important policy choices are based on it, such as banning the communal parties in most of Africa. Because of the multiple confounding factors, it is hard to reach unambiguous conclusions with regard to the explosive potential of sub-national communal identities. A project of the University of Illinois (Social, Political, Economic Event Database, SPEED) may help assess the negative impacts of communal identities as compared to other schisms unavoidably emerging in all societies. SPEED leverages decades of journalist output with AI and human coders; it contains 62,141 newsworthy disruptive events worldwide, from peaceful demonstrations to civil war. It distinguishes between events rooted in class-based conflicts, anti-government sentiment, etc, besides those rooted in socio-cultural animosities. This allows comparing the frequency and intensity of "communal" events to those of events rooted in other cleavages. Analysis shows that SCA events are neither the most frequent nor the most violent, and their relationship with the countries' fractionalization indexes is quite elusive. They are more convincingly related to the countries' population size, and to the desire for political rights. The dataset codes some linkages among events, and there is no evidence of a tendency for SCA-rooted events to escalate. If there is any unusual feature of the SCA events as compared to other categories, it's that their occurrence is less antithetical to the countries' development and democracy levels. Developed democratic countries have managed to subdue actions rooted, for instance, in desires to (unconstitutionally/illegally) retain power, but setting and maintaining standards of communal co-existence is still a work in progress.

Authors: Agnes K. Koos, University of Nevada Reno
Kenneth Keulman, Loyola University New Orleans

Contents

INTRODUCTION

After the end of the Cold War and during the long-lasting European peace, and when it turned out that worldwide, civil wars claimed more victims than inter-state wars, new domains of research cropped up, focused on intra-state conflicts. It seems that in the 1990s two larger thematic clusters started to cohere. The domestic conflict studies were mainly interested in civil wars, imported the methods of the inter-state war studies, invested in databases identifying the conflict's participants, duration, number of battle deaths, and tried to establish the typical patterns of circumstances leading up to war. The other thematic cluster focused on a certain type of domestic conflict, rooted in ethnocultural heterogeneity, and adopted a more grounded-theory-like, process-tracing vision. Both clusters evolved and improved over time, the civil war research incorporated data on lesser conflicts (e.g., HIIK[1]) and a classification of the nature of conflict based on the disputed issues (UCPD[2], PRIO[3]). The study of ethnic conflicts developed databases (MAR[4], EPR[5]) allowing for quantitative research designs to pinpoint causes. Yet research on the impact of ethnocultural heterogeneity has remained somewhat separated from the "general" domestic conflict studies, and we believe that this separation contributed to the persistence of some stereotypes about the former. The main stereotype we want to address here is that ethnocultural heterogeneity is the most consequential source of conflict, leading to the bloodiest civil wars. The reason we have to be concerned about this claim is the policies it inspires. If heterogeneity is a source of conflict *per se*, let alone if it is the main source of domestic conflict, suppression of heterogeneity by any means, including forced assimilation and expulsion/ population exchanges, is justified[6]. If not, then policies that accommodate differences are desirable.

The claims we endorse are that (i) ethnocultural differences are but one cause of domestic conflict among many others, and actually, (ii) only a small proportion of ethnocultural differences has ever led to violent conflicts; (iii) they are very likely to act in conjunction with other causes of conflict; (iv) even when they become a primary factor in delineating the belligerents, some underlying economic, political or socio-cultural inequality of the parties is very likely; (v) on average, they don't lead to more violent conflicts than other causes, such as anti-democratic governing; and (vi) in general, their propensity to escalate is not higher than that of domestic conflicts rooted in other causes.

Of course, we fall short of providing ultimate decisive proofs for all tenets of our vision about domestic heterogeneity in this paper, and concerning some important assertions, we just refer to the pre-existing scholarly literature. Yet we have had the opportunity to look into a dataset that allows for the comparative study of domestic conflicts attributed to a wide array of causes, and which strongly supports the features marked with (i), (iii), (v), and (vi) above. The dataset has been created by the University of Illinois's Social, Political, Economic Event Database (SPEED) project, with a substantial contribution of artificial intelligence. Yet for the important task that created the possibility to compare the number, intensity, and evolution of conflicts rooted in different causes, SPEED had to rely on human coders. This is the reason their dataset takes us from 1946 to 2005 only. The project has mined the

[1] Heidelberg Institute for International Conflict Research (https://hiik.de/?lang=en).
[2] Uppsala Conflict Data Program (https://ucdp.uu.se/).
[3] International Peace Research Institute (https://www.prio.org/).
[4] Minorities at Risk (https://cidcm.umd.edu/research/mar).
[5] Ethnic Power Relations (https://icr.ethz.ch/data/epr/).
[6] To do justice to the proponents of effacing heterogeneity, such as Horowitz 1985, they don't really advocate these aggressive measures to achieve more homogeneity. Yet they suggest institutional designs and laws that prevent even peaceful political mobilization along ethnocultural lines, such as ethnic parties.

media coverage of domestic conflicts for later years, as well, but that work was done exclusively by AI, and the identification of the causes (or of the nature) of conflict was not carried out.

Keeping in the spirit of the SPEED data, we asked ChatGPT about the causes of intra-state conflicts (repeatedly). AI will list at least five groups of factors: (i) political issues (rooted in political disagreements or power struggles between different groups over control of government, resources, or territory, perceived weak governance); (ii) economic issues (such as poverty, inequality, and competition for resources); (iii) ethnic and religious differences (most typically when one group perceives that its interests or identity are threatened by another group); (v) historical grievances or unresolved disputes (such as related to colonialism, land ownership, or past injustices); (vi) external interference (such as supporting one side in a conflict, or steps to destabilize the country)[7]. ChatGPT also emphasizes the interrelatedness of various factors and the unlikeliness of any monocausal explanation. It seems that the human coders of the SPEED dataset followed very similar vision and logic when working out their coding scheme, and they continued to improve it in function of the percentage of cases that could not be neatly placed into their initial rubrics. Importantly, they allowed for the multiple coding (that is, for the multi-causality) of conflict events, and the multiply coded cases outnumber the "monocausal" ones.

SPEED introduces the term "socio-cultural animosities" for the type of conflict more usually called "ethnic", "ethno-cultural", or "communal". A large part of the classic conflict and heterogeneity literature uses "ethnic" largely written as a synonym for "communal", that is, for all boundary-makers that delimit ethnic, linguistic, racial, religious, tribal, caste, and regional groups. The addition of "cultural" (as in "ethno-cultural") explicitly extends ethnic onto linguistic and religious, yet regional seems to fall through the cracks. Ethnicity, in its restricted sense, involves groups that have a consciousness of common ancestry, though their commonality mostly lives in their shared language. Communal groups are long-lasting social phenomena, and every person is born and raised in them. In this sense, communal features are ascriptive, they can be attributed to people even if they make conscious efforts to transcend the boundaries of the group they were born into. Yet in general, people don't try to escape their communal group; on the contrary, they build these ascriptive features into their identity and strive to preserve the well-being of their group(s). Otherwise, they expect society at large to accommodate the particularities of their group(s). We seem to be left with two social engineering solutions: either making societies accommodate all groups equally, or effacing the particularistic (group) identities of all people. The latter solution runs into the hard fact that there is no impartial, above-everybody human essence to take the place of the ousted communal features. Countries need lingua franca-s, if not official languages, agreed-upon national holidays, and principles of family laws, for instance. Practically, most of the time, the communal majorities' language, religion, and culture are adopted as general rules in the country, which embitters minorities. The historical records of accommodation policies are not impeccable, either, but sometimes they lead to the unexpected result of *really* effacing communal identities. For instance, in secular states constitutionalizing the freedom of religion, religious group markers have withered to insignificant salience. Several of Europe's historical minority groups, who are protected by CoE and EU minority rights charters, are shown quite patriotic by opinion surveys, that is, well integrated into their home countries. Some minority groups with

[7] This is the only group of factors not touched upon by the SPEED conflict classification scheme. It can be argued that external factors just piggyback upon pre-exiting internal cleavages. Yet in one respect, the omission of the beyond-the-borders concerns interferes with the clear identification of conflicts. A subset of the events rooted in sociocultural animosities is related to anti-immigration sentiment, yet SPEED does not allow for distinguishing between hostility between long co-existing co-citizen communal groups, and hostility towards newcomers from other countries. We will touch upon the implications in the concluding section of the paper.

widespread autonomy, like French Canadians and Scottish, voted against fully separating from their country. And the European integration project shows that various nations are ready to come together in ethnically heterogeneous polities if they have the certainty of having their interests equitably served.

One of SPEED's contributions to the dispute is the possibility to count the frequency of conflicts rooted in socio-cultural animosities and measure their intensity against the actual ethnic heterogeneity of the countries. In most regions of the world, the frequency and intensity of conflicts did not increase while ethnic fractionalization did. This perspective also highlights the fact that out of so many communal cleavages (e.g., a country with 6 ethnic groups has 5+4+3+2+1 = 15 possibilities of conflict based on socio-cultural differences alone!) a very tiny fraction leads effectively to conflicts only. Compare this with the singular cleavage created by political power, where support for, or opposition to government are the polarized positions, yet this singular cleavage leads to newsworthy manifestations of the dissent so very often! According to SPEED's count, more often than the socio-cultural animosities. Or, in modern societies, sociologists delimit either three large socio-economic status groups (upper, middle, lower) or two classes (capitalists and workers). Yet the number of events classified as resulting from "class conflict" in SPEED comes close to the number of SC animosity-based events.

Obviously, there have always been scholars who advocated on behalf of the relative harmlessness of communal diversity. Fearon & Laitin (1996) highlighted that inter-ethnic peace is the default setting of societies, and conflict is much rarer. Others tried to defend heterogeneity against claims that it hinders economic development and political stability[8], and ethnic mobilization against claims that it exacerbates conflicts (Ishiyama 2009). Since the 2010s, a new wave of scholarship started to produce evidence that communal heterogeneity may have beneficial effects on the economy and politics (Gisselquist et al. 2016, Flesken 2018). Important and influential work has been dedicated to exploring the role of inter-group inequalities in the onset and persistence of ethnic conflicts. The MAR project factored in at least three domains of inequality and oppression of minorities: political, economic, and cultural. The EPR project focused on political inequality, but also showed the role of economic inequality in bringing about domestic conflict. After influential papers by Stewart et al 2008, Cederman et al 2011, and Hillesund et al. 2018, the impact of horizontal (inter-group) political and economic inequality on the likelihood of conflict is well established. The impact of cultural inequality, which occurs when the language, worldview, political ideologies, and/or religious practices of a dominant group are imposed on other groups through laws and the educational system, is scarcely studied, though it may be suspected to explain the "residual" variation of inter-group relations, which are not adequately explained by the political and economic disparities among groups.

The SPEED data does not have built-in measures to compare the political, economic, and cultural statuses of the communal groups involved in conflicts they classified as "socio-cultural animosity," thus we could not venture into a full-fledged test of the role of horizontal inequalities, but SPEED speaks to some other disputed claims about the inter-group animosities.

[8] Since the early 2000s, the literature seems to converge on the idea that diversity comes with both costs and benefits. One of the latest summaries states that “Empirical research working with cross-country data finds a negative, or statistically insignificant, relationship. However, analysis at the city level finds a positive effect of diversity on wages and productivity” (Montalvo & Reynal-Querol 2021).

DATA

Inspired by the evolution of artificial intelligence tools, the University of Illinois launched ambitious projects aimed at scanning and summarizing decades of journalist output. They also publicly shared the fruits of these previously unimaginable endeavors. We have become particularly interested in their datasets on conflict data, but indirectly, also used a different scholarly output of the institute.

More specifically, it is the Cline Center for Advanced Social Research (https://clinecenter.illinois.edu/) to which we owe so much gratitude. First, we mention their Composition of Religious and Ethnic Groups (CREG) Project, which produced a dataset of the different communal groups in 165 countries for the post-WW II era. The Cline Center established the relative sizes of the different groups, and Lenka Drazanova[9] calculated the countries' fractionalization indexes year-by-year. Her dataset is called "Historical Index of Ethnic Fractionalization" (HIEF) data, and uses the same fractionalization index formula as the Soviet ELF, Alesina et al. 2003, and Fearon 2003. Yet older datasets extrapolated the values calculated for a certain year onto all previous and later years, disregarding the slow, but real change of the countries' demographics. The big advantage of HIEF is that it captures these changes, and shows some slight increase in the average fractionalization index over decades. The fractionalization values from this brand-new dataset are highly correlated with older fractionalization measures, which does justice to the work of the CREG project. We added the CREG-Drazanova indexes to our compilation of country features from the Quality of Government (QOG) and Varieties of Democracy (V-Dem) project.

Yet the core data for this paper is the dataset brought about by the Cline Center within the frames of the Social, Political, and Economic Event Database (SPEED) Project[10]. We worked with the dataset called "SPEED Global Random Sample 1945-2005" which was posted to the website on February 15, 2018, under the name "spp_public.xls". It is an event data sheet containing 62,141 newsworthy conflictual episodes from all over the world. The events are classified into four large groups as "political expression", "political attack", "disruptive state actions", and "political reconfigurations".

(i) Political expressions include verbal, written, or broadcast expressions, symbolic actions (passive resistance, boycott, honoring/ disrespecting something), forming associations, workplace actions, and mass demonstrations/ marches.

(ii) The categories included in the notion of political attacks are riot/ brawl; assassination; suicide attack; kidnap/ hostage; execution; other personal attack; other property attack; border incident; and siege/ blockade.

(iii) The list of disruptive state actions runs long, includes 57 items[11], and it is important to note that "failure to act", as well as "[state] service suspensions" are among them. Otherwise, the

[9] https://openhumanitiesdata.metajnl.com/articles/10.5334/johd.16

[10] https://clinecenter.illinois.edu/project/human-loop-event-data-projects/SPEED

[11] The full list is: 1= 'Censorship'; 2= 'Disrupting Electronic Communications'; 3= 'Banning Civil Society Group'; 4= 'Imposing Curfew'; 5= 'Issuing Extraordinary Exec Decree'; 6= 'Declaring State of Emergency'; 7= 'Imposing Martial Law'; 8= 'Dissolving Government'; 9= 'Failure to Convene Legislator'; 10= 'Cancel/Annul Elections'; 11= 'Suspend Constitution'; 12= 'Other Extraordinary State Act'; 13= 'Failure to Act'; 14= 'Warning of Formal Action'; 15= 'Threat to Use Violence'; 16= 'Punitive Discharge of Pub Official'; 17= 'Job Action by Public Employees'; 18= 'Mutiny by Armed Forces'; 19= 'Assembly of Coercive Forces'; 20= 'Mil Conduct of Civilian Functions'; 21= 'Other Minimal State Action'; 22= 'Abuse of Police Powers'; 23= 'Abuse of Legal Discretion; 24= 'Abuse of Judicial Discretion'; 25= 'Other Type of Formal State Act'; 26= 'Facility Closures'; 27= 'Service Suspensions'; 28= 'Restricting Movement/Access'; 29= 'Forced Relocations'; 30= 'Proactive Arrest/Detention'; 31= 'Exile'; 32= 'Trespass on Private Space'; 33= 'Confiscation of Property'; 34= 'Other Coercive State Act'; 35= 'Riot'; 36= 'Brawl';37= 'Assassination'; 38= 'Suicide

list runs from smaller-scale "warning of formal action" and "trespass of private space" to full-blown "suspend constitution" and "cancel/ annul elections".

(iv) "Political reconfigurations" refer to constitutional change, ascension to power and removal from power, relatively rarer events.

All events are assigned a value capturing their intensity as compared to other events in their category. Actually, two larger categories are further divided for calculating their intensity value. SPEED distinguishes between "small-bore political expressions" and "mass expressions". Their intensity scales are different. Disruptive state actions are assigned three different intensity scales by distinguishing between the intensity of state violence, tangible repression, and intangible repression. While state violence is the par excellence example of the regime going rogue, tangible repression is the action of identifiable officeholders who may have gone rogue as individuals, not necessarily involving the rot of the whole regime, and intangible repression is a systemic bias of the regime against some social groups. There are a total of six intensity measures in SPEED, all applicable to their own event type only.

SPEED also invested in creating linkages among events. The column "linked" tells us if an event is related to other events in the dataset. "Link_type" signals whether the connection between events is inclusion or temporal succession. Then there is an indicator containing the identifier of the event that may be regarded as a prequel ("from_eid") and a different column containing the identifier of the event that may be regarded as a sequel ("to_eid"). All these indicators connect only the events in the dataset, and events related to a pre-1946 or post-2005 event, also those that are related to events somehow avoiding being included in SPEED, remained unlinked "orphans"[12].

We have not altered the cells of the SPEED dataset in any way, but added more columns to it. On the one hand, we added some calculated columns, and on the other, we merged in a country-level dataset containing economic, political, and demographic information about the countries where the events took place. We focused on the basic development (GDP, schooling, life expectancy) and democracy indicators, plus included several communal fractionalization indexes, and assessments of the inter-group inequalities from the V-Dem project. For this country-level dataset, our gratitude goes to the Quality of Government and the Varieties of Democracy projects, besides the authors of the HIEF (CREG and Drazanova).

The calculated columns addressed the issue that SPEED's coders used flags in multiple checkboxes for establishing the root causes of conflicts. The original column for "events rooted in socio-cultural animosities", for instance, contains "1" both for cases when "OR_sc_animosity" is the single cause of the conflict and when it acted in conjunction with, say, "OR_class_conflict", and "OR_retribution[13]". We

---

Attack';39= 'Kidnap/Hostage'; 40= 'Execution'; 41= 'Other Personal Attack'; 42= 'Other Property Attack'; 43= 'Border Incident'; 44= 'Siege/Blockade'; 45= 'Other Type of Attack'; 46= 'Attempt Assassination'; 47= 'Attempt Suicide Attack'; 48= 'Attempt Kidnap/Hostage'; 49= 'Attempt Other Personal Attack'; 50= 'Attempt Other Property Attack'; 51= 'Conspiracy Assassination'; 52= 'Conspiracy Suicide Attack'; 53= 'Conspiracy Kidnap/Hostage'; 54= 'Conspiracy Other Personal Attack'; 55= 'Conspiracy Other Property Attack'; 56= 'Other Attempt'; 57= 'Other Conspiracy'.

[12] A distinct indicator called "post-hoc" was meant to flag events deemed to be a "reaction to a prior event", yet only 5039 out of 62141 events were assigned this flag (8%), and the definition in the Codebook contains the restriction that the "destabilizing" event was reaction to some prior event. We considered that we did not have enough information about this variable to use it in our analyses.

[13] We did not change the column names assigned by the SPEED authors, but for easy retrieval of the variables in the softwares we worked with, Excel, SPSS, and Power BI, added a prefix marking the nature of the indicator

created a flag for each of the nine root causes identified by the SPEED coders, which takes the value 1 only when they acted alone (such as "OR_sc_anim.Alone"). Then we extended this further and created an ordinal variable that is 1 when the cause acts in conjunction with other causes and 2 when it acts alone ("OR_sc_anim3"). For the study of the linkages among events, we created calculated tables in Power BI ("FROM" and "TO").

We think that the study of the SPEED data through Power BI dashboards can be a rewarding experience, and posted a copy of it to a website (http://agneskkoos.net/speed-data-dashboard/). Some emphatic results of the data analysis will be addressed here in the next two sections.

MULTIPLE CAUSATION, COMPARATIVE INTENSITY, AND ESCALATION

Table 1 introduces the nine causal categories identified by the SPEED coders and the frequencies with which these origins were attributed to events. The third column shows the number of events that received only one "rooted in" assignment, and finally, the proportion of monocausal events as percentage of the total events belonging to that category is displayed. It is easy to see that multi-causality occurs more frequently than monocausality, and in the case of the events rooted in sc_animosity, the proportion of singular causation is 46.5%. The numbers also show that the most frequent motive of domestic conflicts is discontentment with the government in power. Yet, 21.6% of the 61,141 cases remained unclassified. We may assume that those either did not fit neatly in any of the nine origin categories, or they fit into idiosyncratic large clusters of them. In any case, socio-cultural animosities, on their own, were found to be responsible for 14% of all domestic conflict events between 1945 and 2005. There are no objective criteria to evaluate whether this is trivial or too much, yet it is probably a smaller proportion than expected by those who believe in the superiority of communally homogenous nation-states[14].

Table 1. Distribution of event origins in SPEED

Causes identified by SPEED coders	All Occurences	Monocausal ("Alone") #	Monocausal ("Alone") %
Event rooted in socio-cultural animosities?	18,505	8,608	46.5%
Event rooted in anti-government sentiments?	20,794	9,084	43.7%
Event rooted in desire for political rights?	7,576	1,902	25.1%
Event rooted in class-based conflict?	9,076	4,031	44.4%
Event rooted in ecological resource scarcities	629	186	29.6%
Event rooted in desire to retain political power?	6,215	2,202	35.4%
Event rooted in desire for personal security?	2,663	1,507	56.6%
Event rooted in desire for retribution?	2,682	709	26.4%
Event rooted in desire to maintain public order?	3,540	1,820	51.4%
Total with multiple counts	71,680		
Total unclassified	13,441		
Unduplicated # of events in the dataset	62,141	30,049	48.4%

group. The columns containing root cause data got the prefix "OR", columns on intensity of conflicts got "IY", while those pertinent to event description/classification were marked "ED".

[14] And we cannot help remarking that some of these events rooted in sociocultural animosities are likely to have occurred in quite homogenous countries, as directed against migrants/immigrants.

The "Total with multiple counts" hinges on the way we count them. In accounts that factor in all the co-occurring pairs, it may go up to 85,747. Table 2, which illustrates the co-occurrence of different origins, evidences this counting issue, as well.

Table 2: Co-occurrence of various event origins in SPEED

Causes identified by SPEED coders	OR_sc_animosity	OR_anti_gov_sentiments	OR_pol_desires	OR_class_conflict	OR_eco_scarcity	OR_retain_power	OR_pers_security	OR_retribution	OR_pub_order	Total
OR_sc_animosity		5,286	2,554	1,092	138	1,674	623	942	772	
OR_anti_gov_sentiments	5,286		3,452	3,169	246	1,739	180	856	222	
OR_pol_desires	2,554	3,452		611	46	797	168	290	218	
OR_class_conflict	1,092	3,169	611		169	671	173	279	292	
OR_eco_scarcity	138	246	46	169		6	17	9	10	
OR_retain_power	1,674	1,739	797	671	6		58	332	408	
OR_pers_security	623	180	168	173	17	58		60	171	
OR_retribution	942	856	290	279	9	332	60		119	
OR_pub_order	772	222	218	292	10	408	171	119		
Multiply coded total (# pairs)	13,081	15,150	8,136	6,456	641	5,685	1,450	2,887	2,212	55,698
Monocausal	8,608	9,084	1,902	4,031	186	2,202	1,507	709	1,820	30,049
	21,689	24,234	10,038	10,487	827	7,887	2,957	3,596	4,032	85,747

Despite the multitude of multiple codings, the statistically significant positive correlation coefficients between various origins are rare and may strike us as substantively small[15]. Yet sociocultural animosities demonstrate a clear tendency to co-occur with struggles for political rights and with desires for retribution (Table 3). The Gamma for the pair of OR_sc_animosity and OR_pol_desires is 0.104, and all other tests for ordinal variables (such as Chi-Square, Kendall's Tau-B and C, Somers' d) signal a statistically significant positive relationship between the pair. It is to note that events rooted in the desire for political rights are positively correlated with anti-government events, as well.

Table 3. Correlation matrix of event origins in SPEED[16]

Correlation matrix	Event rooted in socio-cultural animosities?	Event rooted in anti-government sentiments?	Event rooted in class-based conflict?	Event rooted in desire for political rights?	Event rooted in desire to retain political power?	Event rooted in ecological resource scarcities	Event rooted in desire for personal security?	Event rooted in desire to maintain public order?
Event rooted in anti-government sentiments?	-.068**							
Event rooted in class-based conflict?	-.161**	.013**						
Event rooted in desire for political rights?	.032**	.096**	-.069**					
Event rooted in desire to retain political power?	-.021**	-.039**	-.036**					
Event rooted in ecological resource scarcities	-.017**	.012**	.035**	-.015**	-.030**			
Event rooted in desire for personal security?	-.030**	-.120**	-.049**	-.038**	-.055**	-.008*		
Event rooted in desire to maintain public order?	-.043**	-.142**	-.044**	-.045**	.012**	-.018**		
Event rooted in desire for retribution?	.025**		-.025**	-.009*	.017**	-.014**	-.021**	-.012**

Next, Table 4 summarizes the average intensity values of the events by origin type. Since SPEED introduced six measures of conflict intensity, six columns are included, and they show that the events rooted in sc_animosity don't lead any of the six violence scales. (For added visibility, red font shows the values that are above the value for communal issues motivated events.) On two intensity indicators,

[15] Sadly, the binary variables are not really suitable for Pearson correlations, and the large number of cases is also antithetical to having high correlation coefficients. Yet the strength of the association between OR_sc_animosity and OR_pol_desires also varies by politico-geographical regions. It is 0.101 in the Caribbean and 0.083 in 1st World countries, while insignificant in most of Africa and Asia. The next section will dwell longer on these regional, or rather development-related differences.

[16] Here and in all following tables involving correlation coefficients, only the significant ones are displayed. The stars show the significance level of the coefficients, as SPSS routinely reports them.

sc_animosity scores below the overall average, and one of these is the "intensity of political violence", measuring the aggressiveness of the event category "political attack".

Table 4. Average intensity of events by origin and intensity type

Origins	Intensity of small-bore political expression	Intensity of mass expression	Intensity of political violence	Level of intangible repression	Intensity of state violence	Intensity of tangible state repression
OR_sc_animosity.Alone	**1.937**	**3.781**	**2.655**	**2.417**	**6.956**	**2.231**
OR_anti_gov_sentiments.Alone	1.806	2.883	2.743	2.827	11.457	1.890
OR_class_conflict.Alone	1.938	2.858	2.481	2.117	5.194	2.092
OR_pol_desires.Alone	1.769	4.142	2.697	2.010	5.878	2.198
OR_retain_power.Alone	2.329	1.415	2.758	2.275	100.587	2.058
OR_eco_scarcity.Alone	1.987	2.879	1.849	2.067	2.788	2.559
OR_pers_security.Alone	1.589	5.950	2.757	2.315	4.195	2.357
OR_pub_order.Alone	1.482	1.733	2.533	2.328	5.013	2.238
OR_retribution.Alone	1.855	1.292	2.577	2.205	4.263	1.969
Mixed/Unmarked Origin	1.948	4.081	2.700	2.225	4.851	2.209
Average of all events	1.897	3.468	2.685	2.256	11.577	2.195

Yet there are two caveats about the relative harmlessness of communal conflicts as suggested by Table 4. One is about the number of political attacks in each category, and the other is about the temporal trends of intensity.

Table 5 shows the number of political attacks by origin category, and we have to admit that the proportion of political attacks (relative to expression and state action) is the highest in the "OR_sc_animosity.Alone" category (Table 5).

Table 5: Frequency of event types across origin groups

Origin_category	Political Expression	Political Attack	State Disruptive Action	Political Reconfiguration	Unmarked Event Category	Total
OR_sc_animosity.Alone	1310	5172	2034	0	92	8608
OR_anti_gov_sentiments.Alone	4727	3632	371	331	23	9084
OR_class_conflict.Alone	2364	901	744	0	22	4031
OR_pol_desires.Alone	361	905	628	0	8	1902
OR_retain_power.Alone	80	342	1779	0	1	2202
OR_eco_scarcity.Alone	60	72	40	0	14	186
OR_pers_security.Alone	52	70	1315	0	70	1507
OR_pub_order.Alone	51	52	1700	0	17	1820
OR_retribution.Alone	91	206	412	0	0	709
Mixed/ Unmarked Origin	8229	13016	9453	416	978	32092
Total	**17325**	**24368**	**18476**	**747**	**1225**	**62141**

It seems that the paradox of high number of political attacks with relatively low intensity is explained by the type of political attacks that typify the socio-cultural animosity-based conflicts. 65.4% of them fall in the "other personal attack" category, which sounds like a residual category for cases that don't neatly fit into the traditional categories of political attacks on persons, such as assassination and execution.

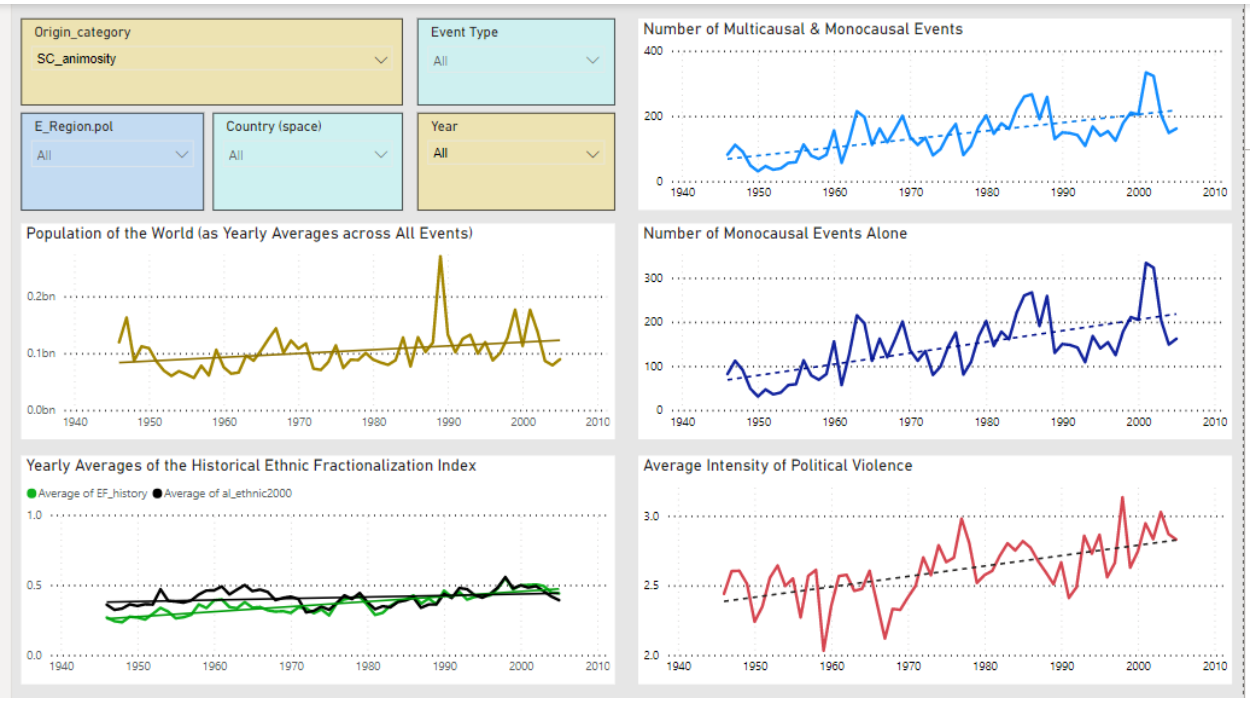
Table 7: Sub-categories of political attack by origin type

Political Attack Type	OR_sc_animosity.Alone	OR_anti_gov_sentiments.Alone	OR_class_conflict.Alone	OR_pol_desires.Alone	OR_retain_power.Alone	OR_eco_scarcity.Alone	OR_pers_security.Alone	OR_pub_order.Alone	OR_retribution.Alone	Mixed/Unmarked Origin	Total	Average Intesity of Type
1="Riot or Brawl"	565	443	126	30	13	44	3	3	19	1350	2596	2.077
2="Assassination"	65	335	12	28	5	0	1	0	10	450	906	3.730
3="Suicide Attack"	126	52	1	13	12	1	1	0	1	240	447	4.008
4="Kidnap/Hostage"	279	215	78	60	22	4	9	2	25	941	1635	3.489
5="Execution"	25	12	10	3	3	0	0	1	5	87	146	3.679
6="Other Personal Attack"	3382	2140	455	548	228	6	37	43	92	7867	14798	2.831
7="Other Property Attack"	676	399	206	209	52	15	10	2	51	1875	3495	1.749
8="Border Incident"	18	13	2	4	4	2	9	1	2	102	157	2.542
9="Siege/Blockade"	9	3	1	0	1	0	0	0	1	30	45	2.069
Unclassified Attack Type	27	20	10	10	2	0	0	0	0	74	143	2.061
Total	5172	3632	901	905	342	72	70	52	206	13016	24368	
Type 6/Total	65.40%	58.92%	50.50%	60.55%	66.67%	8.33%	52.86%	82.69%	44.66%	60.44%	60.70%	

We may wonder which type of event expresses more social tension: a series of mass demonstrations mobilizing thousands, or an assassination carried out by a few dozen people? There is no answer, not even a suggestion to rank them. Further, more verifiably, we may consider the question that which of them is more likely to escalate into a more serious conflict? We attempted to use the SPEED dataset to address this question, but before presenting this test, we want to remark on the temporal trends of the conflict events recorded by SPEED.

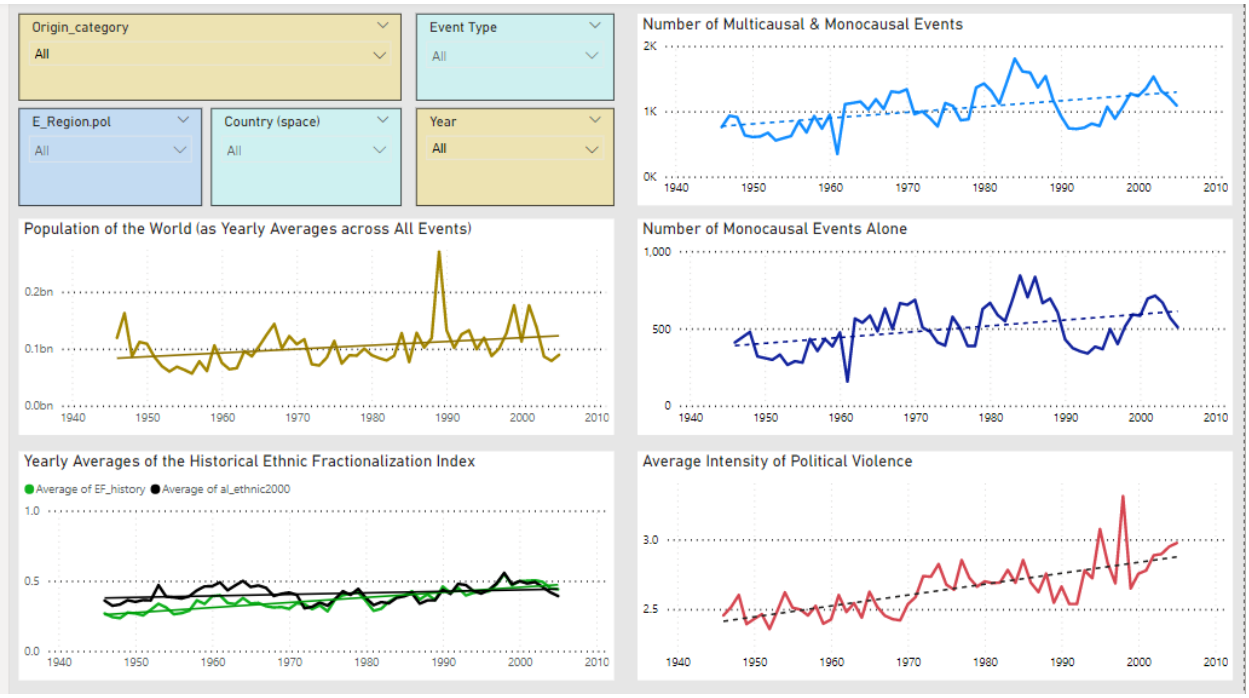
One of the Cline Institute's White Papers on the data and the data dashboard set up by us both show that the number of events rooted in SC_animosity, as well as the intensity of political violence associated with them, increased between 1945 and 2005. The right-side visuals in Chart 1 do not leave any doubt about these trends.

Chart 1: Dashboard screenshot with visuals reporting on events rooted in socio-cultural animosity



Yet the fact is that the trend of all events, independently of their origins, was the same worldwide (Chart 2).

Chart 2: Dashboard screenshot with visuals reporting on all events with all origins



Both the number of events and their intensity went up during the second half of the 20th century. It is hard to find any overarching explanation for the increase in domestic conflicts, and actually, disaggregating the data by event type and political-geographical regions, shows that conflicts may play out very differently in various settings[17]. Here we would like to make just two points and a caveat. (i) The number of conflict events is positively correlated with the population numbers. The population of the globe quasi-doubled between 1945 and 2005. Yet the SPEED data contains a varying number of events in any country for the period, thus the by-year population averages calculated for our visuals are contaminated by this selection bias and show a flatter population increase trendline than legitimately expectable. (ii) Ethnic fractionalization within the countries of the world also increased between 1945 and 2005, and this may be expected to affect the number and intensity of events rooted in socio-cultural animosities. We may suspect that sometimes it's the increase in heterogeneity itself that triggers socio-political conflicts. Think about some of the typical ways in which a country's heterogeneity increases. Practically mono-ethnic developed countries have received large influxes of immigrants. There are fertility rate disparities among communal groups, and often, the historically marginalized minorities grow fast enough to threaten the grip on power of the establishment groups. Some cultural groups that seemed to have completely assimilated to a dominant culture, "awaken", and fight for their own identity and cultural rights, like the Scots, Bretons, and Silesians. When displaying the increase of fractionalization, we run into the same selection bias issue as in the case of population increase. A trendline drawn by factoring in each country only once a year is steeper than we have in our dashboard. For the sake of comparison, we also charted in the visual Alesina et al.'s 2003 ethnic fractionalization

[17] For instance, as the next section will detail, class conflicts trend downward in general, and political attacks trend downwards or stagnate in developed countries.

index, which extrapolates the year 2000 ethnic composition for all years before and after 2000. Visibly, that line also shows some oscillation, though, in a one-country-by-year arrangement, that should be a straight line. It's reassuring to see that the two lines co-vary (they are highly correlated), yet the green line, which charts the historical fractionalization values, has an explicit increasing tendency, at least on global scale (some regions are exceptions). (iii) The caveat we want to add is that the increase in the number of events may be, to some extent, an artifact of the journalist activity that was captured by the SPEED data. As the resources for, and the techniques to collect and forward information about events in remote places have improved, we can expect more and more events to break through and make headlines. There are no known ways to prove or refute this assumption, let alone a way to gauge the extent of this possible bias. Yet this "artifact effect" is plausibly impacting the number of events only, and much less the intensity measures.

The final test in this section is pertinent to the linkages among events. We have to bear in mind that the temporal limits of the dataset truncate some series of events, that is, a number of events may show up as "orphans" because their prequel or sequel remained outside the dataset's timeframes. We checked on the propensity of conflict events to escalate by comparing the intensity of "prequels" and "sequels". It is to note that SPEED avoids assigning causal significance to temporal relationships, they don't use the prequel and sequel labels. And we also use them without postulating that there is a direct causal impact from the prequel (in SPEED's terminology: a "from" event) to the sequel (a "to" event)[18]. Yet we believe that some type of complex social causation between related events is highly likely.

The literature often expects social conflicts, and mainly conflicts of communal nature, to easily spiral out of control and lead to more and more violent iterations of the collisions. The SPEED data does not confirm these expectations for the intensity of political violence measure, which mainly covers the political attack event type. Sequel political attacks do not show up as more violent than prequels either in the whole sample (Chart 4) or in the case of events rooted in socio-cultural animosity (Chart 3)[19].

Intensities of other types than that of political violence show a more variegated picture, but no definite trend toward escalation is transpiring from the data. For instance, Chart 3 shows the intensity of state violence as increasing from 4.55 (prequels) to 4.73 (sequels) and then to 4.95 (second order sequels, "to-to"-s. Yet specifically for the communal conflicts, state violence subsides in the sequels, while state repression shows an increasing trend in the second-order sequels.

[18] We may imagine various constellations of inherently related events that do not cause each other, but are induced and shaped by the same field of forces within a society. For instance, employees of Starbucks win union contracts in a number of units in the US, and a few months later, the Amazon employees also succeed in a number of warehouses.

[19] The only origin type where the average sequel is more violent than the average prequel is personal security, yet the number of sequels in this category is only 59.

Chart 3: Dashboard screenshot with visuals reporting on intensities of prequels and sequels – Events rooted in SC animosity

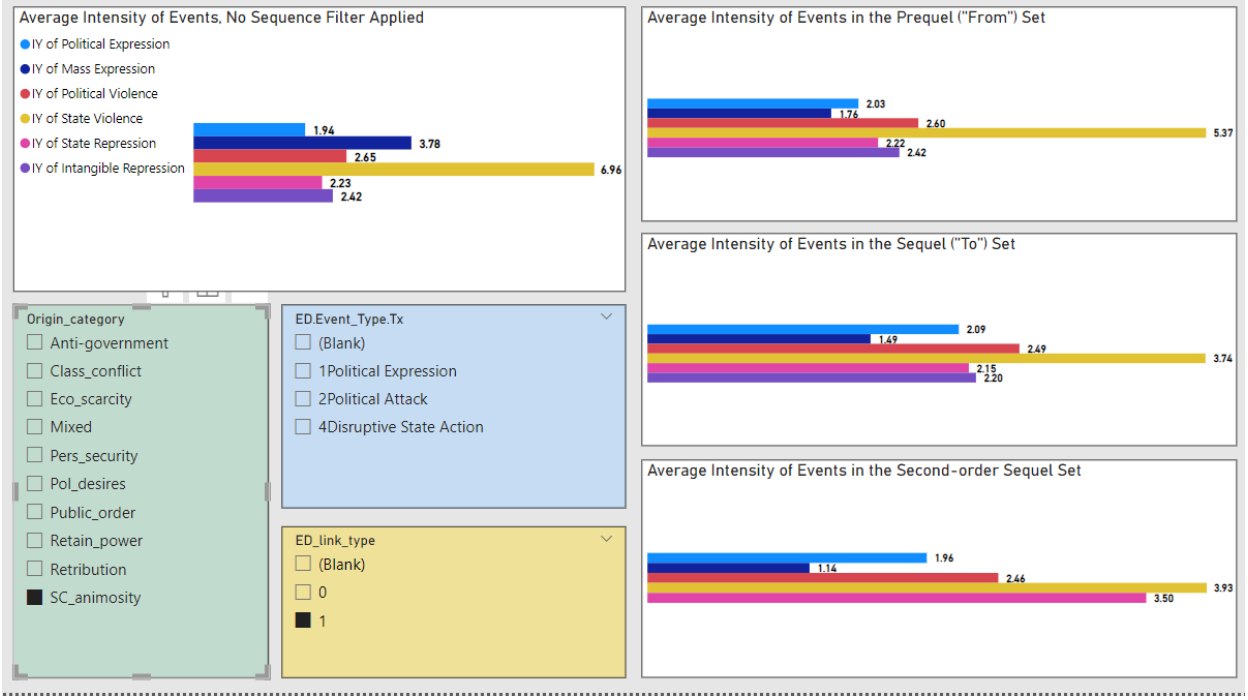
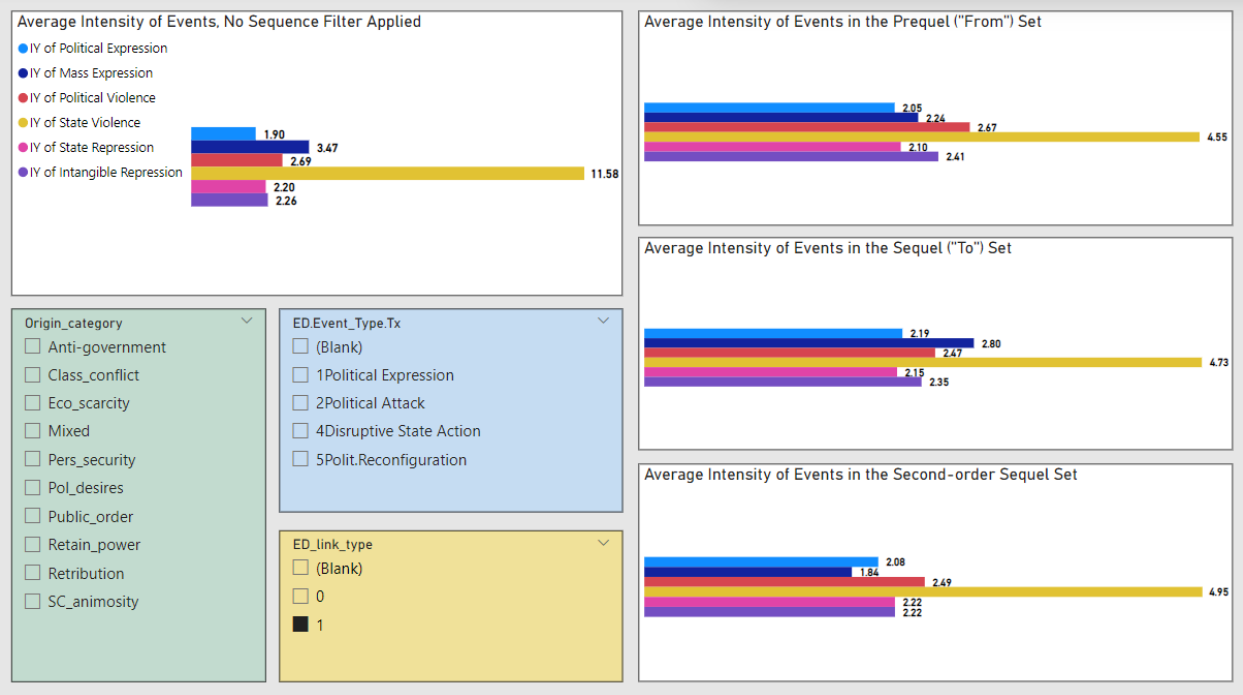


Chart 4: Dashboard screenshot with visuals reporting on intensities of prequels and sequels – All events



THE BACKGROUNDS: DEVELOPMENT, DEMOCRACY, EQUALITY, HISTORY

Charts 1 and 2 suggested an impact of population size on the number and intensity of conflicts, plus communal fractionalization's impact on the number and intensity of events rooted in sociocultural animosities. Since the SPEED data is a selection on the dependent variable – only conflict events are included, no negative cases of "when a conflict did not occur" – we could not test the impact of the population size by correlating the occurrence of SPEED events with the population size. We entered some other conflict measures in the dataset, and it is the UCDP Type 3 (internal) conflict measure that displays a +0.167 coefficient with the population size, corroborating the commonsense expectation that among more people, more conflicts are possible.

As for the SPEED data itself, we looked at the type of conflicts that are more likely with increasing population, and, indeed, it is only the political expression event type that proliferates when countries are more populous. Political attacks and state disruptive actions don't proliferate. This trend is confirmed by another "foreign" conflict indicator merged into the dataset. The V-Dem project's "Mass mobilization" is positively correlated with population size, while their "Political violence" measure is negatively correlated.

We may suspect that there is some other force at action here, and indeed, further correlations show that the population size and the level of development are positively correlated in this dataset (+0.108). Further, the more developed a country, the more likely it experiences events of political expression, yet the less likely that it experiences disruptive state actions. These relationships are stable across at least nine development measures involving GDP per capita, life expectancy, and education. And they hold across dozens of democracy indicators and equality and fairness measures imported from the V-Dem project. Finally, the regional distribution of the events also shows more political expression events in developed countries, and more state disruptive actions in the developing world[20].

The SPEED dataset does not tell the story of a world becoming gradually more peaceful, where domestic conflicts wither away with development and democratizing. It suggests that (i) the nature and origin of newsworthy expressions of social tensions change over time, and (ii) the state's intervention in these tensions is less and less likely to be biased and violent. Otherwise, development, as we know it, meaning mostly the increase of GDP per capita, and realized mainly through capitalist global arrangements, does not obliterate the accumulation of social tensions that emerge as newsworthy conflict events. Yet democracy, as we know it, rooted in free and fair elections, may allow for the conflicts to be played out without violence. If the number and intensity of political expression and mass expression increase, yet the number and intensity of state violence do not, then some demands of the protesters/demonstrators are mainstreamed and policies are changed to accommodate the requests. Virtuous circles of state responsiveness to emerging social tensions are possible. And a reading of the SPEED data supports their existence.

[20] This constellation is visualized in the dashboard's "By Regions" page. It is to note that there is a "Cambodia85" slicer on that page, because of a probably erroneous entry of an enormous outlier value, which dwarfs all other values in the visual. ("Intensity of state violence" takes the value of 49,774.9, seemingly because of a value of 5.5 million victims attributed to state violence in that year. However, in 1985 control of Cambodia was divided between the Vietnam-supported government and Khmer Rouge rulers. 5.5 million may be the total number of victims of the Khmer Rouge rule in the 1970s plus of the following civil strife during the 1980s. For comparison, state violence during the Chinese Civil War was assigned the intensity value of 1,587.8, for killing 175,000 people in 1949.)

Unfortunately, no convincing regression models can be built with the SPEED data enriched with country-level indicators. We cannot meaningfully ask what country features make conflicts more likely because all cases are conflict events. We may ask whether some types of events are more likely given certain country features. We constructed regression models with the events rooted in socio-cultural animosities as the dependent variable. The models are weak, explaining not more than 4% of the variation, as the Nagelkerke Pseudo-R2 indicates in our ordinal logit models[21]. When trying to disaggregate by event type (political expression, political attack, and state disruptive action), we obtained similarly low model strength values for all of them. In a next step, we ran the same models for the other eight origin types, as well, and only the model on class conflict was somewhat stronger than the model attempting to explain the incidence of events rooted in sociocultural animosities. It seems that the basic social constellations (such as development level, democracy, vertical and horizontal egalitarian arrangements) only slightly influence, but don't predetermine the type of conflict that emerges. That is, social complexity makes up for a highly polyvalent causal substratum. The small model strength values don't allow for a strong claim about the two most surprising findings summarized in Table 8.

First, events rooted in sociocultural animosities are more likely to occur in developed countries than in the less developed or developing. (In Table 8, the red color highlights the significant positive regression coefficients that we obtained instead of the expected negative coefficients.)

Second, the incidence of events rooted in socio-cultural animosities is not impacted by the countries' fractionalization. Strangely, it is the incidence of events rooted in desires for political rights and in maintaining public order, which goes up when fractionalization is higher. On the downside, class conflicts are shown to be less likely to occur when fractionalization is high. Yet we also have to bear in mind that less developed countries are typified by more communal heterogeneity. Events rooted in class conflicts, like those rooted in sociocultural animosities, are more frequent in the developed world. Yet the incidence of class conflicts started to decline during the last decades covered by the SPEED data, while the same did not happen in the case of communal animosities.

[21] In these trials we used the 3-value version of the Origin indicators, which assigned 2 to mono-causal, and 1 to mixed-root occurrences of the origin type.

Table 8: Ordinal logit models on the nine Origin Types

Ordinal Logit Models	SC_anim	Anti_gov	Class_cnfl	Pol_desir	Ret_power	Eco_scrcty	Pers_scrty	Pub_order	Retribut	Expectation
	Estimate									
GDP per capita	0.027		-0.029		-0.008	-0.029	-0.009		0.012	
Electoral democracy index			1.025	0.506	-0.722	0.824	-0.491	-0.561		reduces occurrence (- sign)
Egalitarian component idx ordinal	0.132		-1.008	0.371	0.330	-1.106			0.843	
Physical violence idx ordinal	0.336		1.092	-0.244	-0.446	0.782		0.228	-0.814	
Political polarization (4=deep)	-0.031		0.112	-0.036	0.057	0.191	-0.086	0.071	0.049	increases occurrence (+ sign)
Political violence (4=often)	0.234	-0.023	-0.200	0.127		-0.317	0.041			
Historical fractionalization idx			-0.789	0.373				0.296	-0.530	impacts only SC_anim
	Significance of Covariates (P-values)									
GDP per capita	0.000	0.737	0.000	0.174	0.000	0.000	0.006	0.825	0.000	
Electoral democracy index	0.790	0.181	0.000	0.000	0.000	0.010	0.002	0.000	0.723	
Egalitarian component idx ordinal	0.008	0.276	0.000	0.000	0.000	0.000	0.726	0.407	0.000	
Physical violence idx ordinal	0.000	0.190	0.000	0.001	0.000	0.000	0.503	0.023	0.000	
Political polarization (4=deep)	0.004	0.430	0.000	0.013	0.000	0.000	0.000	0.000	0.031	
Political violence (4=often)	0.000	0.010	0.000	0.000	0.167	0.000	0.054	0.577	0.109	
Historical fractionalization idx	0.109	0.248	0.000	0.000	0.996	0.141	0.342	0.000	0.000	
	Model Strength									
Cox and Snell Pseudo-R2	0.034	0.001	0.029	0.004	0.014	0.002	0.002	0.002	0.003	
Nagelkerke Pseudo-R2	0.043	0.001	0.045	0.007	0.026	0.015	0.005	0.004	0.008	
McFadden Pseudo-R2	0.022	0.000	0.029	0.005	0.018	0.014	0.004	0.003	0.007	
Valid cases	53798	53798	53798	53798	53798	53798	53798	53798	53798	

We also endeavored to acquire more information about the relationships of various types of events with country features through principal component analysis. Typically, 7 factors were needed to explain two-thirds of the variance, yet the major problem was that a strong first PC siphoned up all positive country features (development, democracy, vertical and horizontal equality), and a half dozen weak PCs spelled out a few relationships among the event types and events of various origins. Unfortunately, SPEED itself does not contain measures that could be used as explanatory variables. Thus the explanatory models only use country-level variables, which also prevents multi-level modeling.

All together, correlation matrices, regressions, PCs, and the Power BI visuals, suggest that two Origin types, class conflict, and sociocultural animosities, are more likely to occur in countries that are more developed, more democratic, and more egalitarian, than the countries where events rooted in retaining power and in fear for personal security are more frequent. That is, data show some polarization of the origin types along development, democracy, and egalitarianism measures. Of the two occurring on higher development level, class conflict (as defined by SPEED) is expected to occur in countries with capitalist relations, and the West European evolution suggests that it is more violent in an earlier phase of unbridled capitalism, such as in 19th century Europe, and as the proletariat gets its ways in politics and policies, the class confrontation becomes a parliamentary struggle between Left and Right, with a considerable welfare state in place. Yet with sociocultural animosities, the expectation of most of the literature is that they occur on lower levels of social development (such as pre-national stage for Horowitzians, and pre-post-materialist stage for Inglehartians). Anticipating some theoretical resistance to this finding from the SPEED data, we performed one more data transformation to address the missing "non-event" problem.

This transformation consisted of calculating the yearly averages of event occurrence from the SPEED data. We used the nine "origin" indicators that took value 2 when an event was attributed to only one cause; 1, when the event was attributed to two or more causes, and 0 when the event was unrelated to the respective cause. Power BI readily calculated the yearly averages by country-year, and the

spreadsheet with these nine measures, plus the yearly maximums of the six intensity indicators was merged with our country-year level compilation of QOG, V-Dem, and Historical Fractionalization data, cut back to 1946-2005. In principle, the merged data allowed us to test the impact of country features on non-truncated dependent variables, which covered the possibility that no event happened in a certain year. Yet the merger resulted in a very large number of no-events (that is, blanks, which were transformed into zeroes) in the event occurrence and maximum intensity columns. The country-year SPEED export has had 4,938 rows, while the merged 1946-2005 QOG-V-Dem-HF has had 11,897. No-event country-years vastly outnumber event-filled country-years by about 7 to 5.[22] This is good news for real life, but bad news for research. The distribution of the SPEED variables in their new context is so heavily right-skewed, that no transformation can make them resembling a normal distribution. We still ran several correlations, group means analysis, and even regressions, to compare the results with what we previously found.

*Table 9: Correlations of the SPEED indicators when embedded in country-year dataset**

	OR_anti_gov_AV0	OR_class_confl_AV0	OR_eco_scrcty_AV0	OR_pers_secrty_AV0	OR_pol_desir_AV0	OR_pub_order_AV0	OR_ret_power_AV0	OR_retribut_AV0	OR_sc_anim_AV0
OR_class_confl_AV0	.210**								
OR_eco_scrcty_AV0	.073**	.072**							
OR_pers_secrty_AV0	.094**	.056**	.026**						
OR_pol_desir_AV0	.250**	.074**	.019*	.077**					
OR_pub_order_AV0	.148**	.089**	.033**	.078**	.108**				
OR_ret_power_AV0	.198**	.087**	.027**	.056**	.135**	.130**			
OR_retribut_AV0	.122**	.053**		.057**	.075**	.073**	.109**		
OR_sc_anim_AV0	.195**	.075**	.038**	.119**	.193**	.118**	.124**	.119**	
IY_MX_intang_rep0	.250**	.181**	.042**	.134**	.176**	.259**	.307**	.156**	.215**
IY_MX_mass_express0	.383**	.368**	.079**	.095**	.156**	.148**	.121**	.084**	.233**
IY_MX_pol_express0	.474**	.264**	.094**	.149**	.238**	.191**	.190**	.147**	.320**
IY_MX_pol_viol0	.525**	.276**	.104**	.213**	.366**	.225**	.263**	.207**	.494**
IY_MX_st_repress0	.316**	.240**	.071**	.242**	.217**	.345**	.314**	.198**	.324**
IY_MX_stat_viol0	.271**	.152**	.061**	.253**	.233**	.206**	.275**	.158**	.326**

**The log 10 transformations of the variables were used to produce the matrix.*

Table 9 shows a correlation matrix with positive values only. The huge number of zeroes from the non-event country-years pushed the relationships from negatively correlated to positively correlated. Yet the magnitude of coefficients (emphasized with green highlights) reinforces the previously detected connection between events rooted in sociocultural animosities and in struggle for political rights (+0.193**). In the upper part of the matrix, only coefficients of the anti-government movements are stronger than this[23]. The lower part shows the relationships between origin and conflict intensity, and the strongest correlations are in the columns of anti-government movements and sociocultural animosities.

Further correlations, between the SPEED variables and the country features also confirmed the previous findings. Larger population means more conflict, of every type, and the intensity of the conflicts may

[22] Numbers were influenced by anomalies such as SPEED events recorded without territorial localization, and events recorded in non-sovereign territories that do not have country features in QOG-V-Dem-HF for the respective year (e.g. Germany before 1949, colonies before independence, Puerto Rico, Vatican City etc.).
[23] The strong correlations between anti-government movements and almost all other origin types suggests that people do not fight government only because it is a bad government (e.g, corrupt, inefficient), but also because it promotes policies that are disadvantageous to some groups (such as workers, or powerless communal groups).

also be higher. Communal fractionalization is negatively related to the likelihood of communal animosities, and the intensity of most types of conflicts is lower when the fractionalization index is larger. For buttressing the robustness of the findings, we involved Alesina's fractionalization indicators, as well, and conflict indicators from other datasets[24].

Table 10: Correlations of the SPEED indicators with country features & controls, when embedded in country-year dataset

	Population	Historical fractionalization idx	Ethnic Fractionalization (Alesina)	Language Fractionalization (Alesina)	V-Dem Mass mobilization	UCDP Armed conflict, internal	Civil war
OR_anti_gov_AV0	.130**	-.054**	-.032**	-.074**	.229**	.159**	.107**
OR_class_confl_AV0	.081**	-.082**	-.085**	-.116**	.133**	.044**	
OR_eco_scrcty_AV0	.049**			-.022*	.044**	.026*	
OR_pers_secrty_AV0	.070**				.073**	.097**	.077**
OR_pol_desir_AV0	.076**				.130**	.138**	.141**
OR_pub_order_AV0	.082**				.114**	.076**	.058**
OR_ret_power_AV0	.088**				.113**	.141**	.125**
OR_retribut_AV0	.048**	-.034**	-.032**	-.037**	.059**	.056**	.035**
OR_sc_anim_AV0	.153**	-.059**	-.032**		.161**	.189**	.106**
IY_MX_intang_rep0	.249**		-.024*	-.036**	.213**	.164**	.093**
IY_MX_mass_express0	.240**	-.098**	-.103**	-.108**	.278**	.102**	.030**
IY_MX_pol_express0	.274**	-.096**	-.088**	-.100**	.295**	.197**	.104**
IY_MX_pol_viol0	.222**	-.037**	-.023*	-.054**	.293**	.320**	.275**
IY_MX_st_repress0	.260**	-.042**	-.040**	-.051**	.266**	.221**	.145**
IY_MX_stat_viol0	.204**	-.034**			.226**	.332**	.295**

Yet for practical purposes, we want to learn about the relationships of the conflicts with country features over which citizens have some control, or, at least, can tell whether development will bring a more peaceful world or not. Thus, we are interested in the relationships of the conflict indicators with features such as economic development, democracy, good governance, equality, and inclusion. Table 11 contains seven positively coded variables, where the expectation is a negative coefficient sign, and four negatively coded governance variables, where the expectation is a positive sign. Yet the table features a great deal of red color, with which we marked the unexpected direction of the relationships.

[24] In Table 10, too, the Log10 transformation of the SPEED variables were used.

Table 11: Correlations of the SPEED indicators with country features

	GDP per capita --- estimate	Electoral democracy index	Unified democracy score Mean	Unified democracy score Median	Physical violence index ordinal	Regime corruption	Political corruption index	Neopatrimonial Rule Index	Clientelism Index	Egalitarian component index ordinal	Equal distribution of resources idx ord.
OR_anti_gov_AV0		.063**	.031**	.031**	-.045**	.043**	.053**			.048**	.026*
OR_class_confl_AV0	.022*	.137**	.137**	.136**	.085**	-.104**	-.103**	-.135**	-.127**	.082**	.070**
OR_eco_scrcty_AV0						.027**	.032**				
OR_pers_secrty_AV0	-.022*	-.021*	-.033**	-.033**	-.059**	.041**	.045**	.042**	.030**	-.021*	-.022*
OR_pol_desir_AV0			-.035**	-.035**	-.095**	.030**	.038**	.035**		-.025*	-.037**
OR_pub_order_AV0	-.035**	-.025*	-.044**	-.044**	-.057**	.049**	.056**	.042**			-.023*
OR_ret_power_AV0	-.064**	-.087**	-.119**	-.120**	-.140**	.093**	.096**	.109**	.052**	-.051**	-.053**
OR_retribut_AV0	.027**				-.039**					.026*	.036**
OR_sc_anim_AV0	.045**	.058**	.042**	.042**	-.032**	-.023*		-.050**	-.077**	.059**	.046**
IY_MX_intang_rep0		-.035**	-.067**	-.067**	-.092**	.025*	.032**	.040**	-.036**		
IY_MX_mass_express0	.092**	.185**	.168**	.167**	.100**	-.115**	-.109**	-.157**	-.162**	.136**	.108**
IY_MX_pol_express0	.053**	.112**	.087**	.087**		-.063**	-.052**	-.087**	-.127**	.097**	.076**
IY_MX_pol_viol0		.066**	.040**	.039**	-.090**	.049**	.065**				-.022*
IY_MX_st_repress0		-.021*	-.049**	-.050**	-.109**	.037**	.046**	.041**	-.036**		
IY_MX_stat_viol0	-.060**	-.051**	-.091**	-.091**	-.150**	.099**	.107**	.089**	.047**	-.045**	-.066**

As previous findings also suggested, the political expression event type (of which intensity coding is "intensity of mass expression" and "intensity of political expression"), is more prevalent in developed, democratic, and even more egalitarian settings. It is the state disruptive action event type that tends to subside with development and democratization. Political violence type events seem to be not, or ambiguously related to these basic country features. And the finding about a polarization of the events with different origins is also supported by Table 11. Class conflict does definitely occur in more developed, more democratic, and more egalitarian settings, and communal animosities show a similar tendency. Anti-government movements and events rooted in retribution are also little mitigated by positive country features.

Finally, we replicated the regression models in Table 8, this time in OLS format. Unfortunately, these models are also very weak[25], and one even failed to show up as significant[26]. Only two origin types seem to be unambiguously deterred by the development/democracy/egalitarianism triad: the desire to retain power and fear for personal security. All others relate ambiguously to these country features.

[25] By adding Population size to the models, which has a significant positive correlation with the DVs, their strength increases about 1 point, that is, population explains a further 1% of the variation.
[26] It was exactly at the boundary with a significance of 0.050.

Table 12: Regression models replicating the Table 8 models

OLS Models, with Log10 transformation of the DVs	OR_sc_anim_AV0	OR_anti_gov_AV0	OR_class_confl_AV0	OR_pol_desir_AV0	OR_ret_power_AV0	OR_eco_scrcty_AV0	OR_pers_secrty_AV0	OR_pub_order_AV0	OR_retribut_AV0
	Standardized coefficient (Beta)								
GDP per capita --- estimate	0.100	0.052		0.048					0.064
Electoral democracy index	0.046	0.067	0.127		-0.065			-0.055	
Egalitarian component index ordinal	0.105	0.063		0.048				0.047	0.060
Physical violence index ordinal	-0.076	-0.062	0.067	-0.062	-0.074				-0.073
Political polarization			0.074						0.036
Political violence	0.190	0.237	0.077	0.209	0.122		0.116	0.124	0.056
Historical fractionalization idx	-0.063	-0.075	-0.063		-0.027		-0.041		-0.041
	Significance								
GDP per capita --- estimate	0.000	0.000	0.930	0.000	0.252		0.153	0.522	0.000
Electoral democracy index	0.018	0.000	0.000	0.200	0.001		0.605	0.005	0.789
Egalitarian component idx ord.	0.000	0.000	0.138	0.004	0.041		0.820	0.005	0.000
Physical violence index ordinal	0.000	0.002	0.001	0.002	0.000		0.071	0.309	0.000
Political polarization	0.059	0.211	0.000	0.991	0.083		0.059	0.014	0.024
Political violence	0.000	0.000	0.000	0.000	0.000		0.000	0.000	0.001
Historical fractionalization idx	0.000	0.000	0.000	0.088	0.028		0.001	0.079	0.001
	Model strength								
Adjusted R2	0.044	0.056	0.031	0.040	0.039	nsig	0.012	0.018	0.012
Number of cases	7317	7317	7317	7317	7317	7317	7317	7317	7317

The two main changes, as compared to Table 8, are related to the first two origin types. Communal conflicts' relationship with the fractionalization index emerges as negative. And the anti-government movements turn up as positively related to GDP, democracy, and egalitarianism, though the impact of the "positive triad" is smaller than in the case of the events originating in sociocultural animosities.

DISCUSSION AND CONCLUSIONS

The SPEED dataset has offered a unique possibility to study domestic conflicts, with several exceptional merits. First, it is the scope of the enterprise, covering the whole globe, and for a six-decade stretch from 1946 to 2005. Second, the authors coded the "roots" of the events, that is, performed a classification that borders on establishing the causes of conflicts. Third, they also performed a classification of the event types that distinguished between violent and non-violent actions. Fourth, they distinguished between actions initiated by persons operating the state and actions by non-state actors. And fifth, conflicts of various intensity were considered, and intensity scores assigned, which in longitudinal panel data allows for process tracing studies.

We tried to make good use of this wealth of data in order to address the issue of comparative danger of different kinds of domestic conflict, and also tried to contextualize their occurrence and intensity by adding country features to the SPEED dataset. The SPEED dataset was meant to thoroughly describe the conflict events, but it does not contain country (or society) features that in traditional causal analyses are probed as determinants of conflicts, such as communal fractionalization, socio-political exclusion of certain groups, surplus of unemployed young men, to name a few. The origin types ("event rooted in") were established primarily based on the features of the perpetrator and victim groups, for instance, a conflict between Hindu and Muslim groups in India typically showing up as event rooted in socio-cultural animosities.

We think that "animosity" may be conceived of as a long-term attitude, reaching various degrees of relevance and intensity, waxing and waning in function of social contexts. Moreover, positive or negative emotions toward other groups, as well as the very conceptualization of "we-group" and "other group(s)", may also be considered social facts, brought about by other social facts. Communal groups are relatively long-lasting creations of history, geography, cultural contacts, and economic development, for instance, capitalist production's needs are largely regarded as nation-forming forces, which broke down the boundaries of feudal entities, promoted unitary language, education, and secular loyalty to the nation-state. Capitalism also brought about proletarian class consciousness (the working class's "we-group" as opposed to the "owner group/class"), and class conflicts. The parallel intends to highlight that group consciousness forms based on the perceived differences between groups, but it becomes a salient feature of people's worldview and a source of inimical feelings when a group develops the creed that the other wronged them or competes with them for scarce resources. Those who advocate the effacement of sub-national communal identities because they are inherently dangerous, basically claim that animosity between communal groups is a persistent state of affairs, and unscrupulous political entrepreneurs may easily turn the latent hostility into violent conflict. As a matter of fact, unscrupulous political entrepreneurs typically try to make a communal group believe that they were wronged and need to revenge, or that they have to prevent the other group from benefitting from unfair advantages in some competition. The big question is whether a not wronged and not-at-unfair-competition group would follow these false prophets into violent conflicts or not. Unfortunately, society has always been filled with frictions, and most social groups, delimited by either communal, socioeconomic, political, gender, or sexual orientation features, have something in their past to complain of and for which to blame other groups. Even the privileged ones may complain of harsh social justice events such as the French aristocracy of Jacobin dictatorship, and former owners when communist/socialist governments nationalize their property. Reasons for group enmity abound, and in general it's not the dearth of political entrepreneurs that prevents them to bring about violent group conflicts. We can see two plausible explanations that there are so many "no-event" years and that many societies may go long stretches of time without significant group violence.

First, people are not unidimensional, fixated on one identity. They are, and most perceive themselves as being, at the intersection of several social identities. They belong to several groups, and typically, not all of those are on the trailing end. (Muslim in a Christian state, but wealthy; fresh immigrant, but accomplished professional; woman, but White.) When people have the choice to construct their identity, they strive to address and include all their group belongings. Unilateral focus on any identity comes only under duress, when the state or other influential groups confine some people to a certain identity (as in Nazism, and apartheid), or when members of a group suffer systematic disadvantages as members of a certain group (proletariat, "ranked" communal groups).

Second, most societies allow for some non-violent forms of conflict resolution, and for forms that don't count as newsworthy "events" at all. There are court cases against hate speech and harassment; elections to oust clientelistic officeholders; and trade unions to negotiate work contracts. By definition, democratic institutions are set up as arenas for articulating a common will and carrying out disagreements among members of the demos in regulated ways[27].

[27] Unfortunately, older, classically liberal constitutions, such as that of the US, only factor in disagreements among individuals, which are attributed to their upbringing, knowledge, and decision-making abilities. The Framers did not consider systematic, interest-driven conflicts among groups of people, neither class-based or ethnic. This is why they did not regulate political parties, for instance.

This perspective converges with the finding from the SPEED data that neither sociocultural fractionalization, nor class segmentation (inequality and exclusion) are good predictors of conflict events. Moreover, it can be claimed that sociocultural animosity is not a sufficient cause of communal conflicts. In general, SPEED's "origin types" typify the nature, rather than the causes of the events. A large body of literature claims that the causes of domestic conflicts lie in the inequalities and injustices plaguing societies, besides the dishonest and greedy behavior of the persons operating the state. Another stream of literature looks at the opportunities of various groups to organize and carry out violent actions, and we have to admit that large amounts of oppression and injustice may go without noticeable resistance when the victims are too powerless to make their voice heard. If we accept that visible conflict is preferable to the silence of powerlessness, the historical trajectories of conflict occurrence and intensity revealed by the SPEED data may allow for an optimistic interpretation.

In a first reading, it sounds disappointing that the social constellations we reasonably expect to mitigate conflicts, economic development, democracy, and egalitarianism, have only moderate, partial impact on certain types of conflict. But there are some promising trends, such as:

- State disrupting behavior lessens in economically developed and democratic states;
- Events rooted in the desire to illegitimately retain power and those rooted in fear for personal security also lessen in economically developed and democratic states;
- The event type of which occurrence is positively related to development level is political expression, basically non-violent exercise of democratic rights.

Undeniably, the SPEED data shows that class conflict and conflicts rooted in sociocultural animosities typify developed societies, rather than the less developed. Of the two, class conflict seems to have matured in a declining pattern by 2005. The word "matured" is meant to convey the belief that in a sense, the labor movement has become victorious in many parts of the world. The Soviet bloc was defeated, but the worker-friendly welfare state did increasingly become the norm in the 20th century. (Unfortunately, reversions also happened, as in the US.) In parallel with codifying some workers' rights, the ways and methods through which class conflicts can be played out were also institutionalized. For instance, collective bargaining is highly regulated, and most of the time the negotiations remain the organizations' internal affairs. It's only when the bargaining breaks down, and the workers go on strike, that the conflict becomes a newsworthy event.

The pattern of class conflicts being associated with most developed countries and starting to decline at the end of the 20th century can be theorized as a successful movement's changing gear and methods. Labor movements preceded the SPEED dataset, and indeed, by 1946, they achieved very serious socio-political changes in the most developed countries: full suffrage, shorter workday, workplace safety regulation, and strong leftist political parties that worked on implementing the welfare state with free education and universal healthcare. (Even the US went in that direction during the New Deal and possbly under Kennedy.) Exactly like the intra-organization bargainings, the everyday political activity in legislatures and governments of all levels are not newsworthy events.

This seems to be the vantage point from where the efforts to suppress communal organizing (such as bans on ethnic and religious parties) look as being on the wrong side of history. Western evolution enshrined freedom and equality, and there are no convincing arguments for denying any group the right to pursue these. Events rooted in sociocultural animosities are associated with struggles for political rights, and also motivated by experienced social injustices. Since these are social issues that need to be

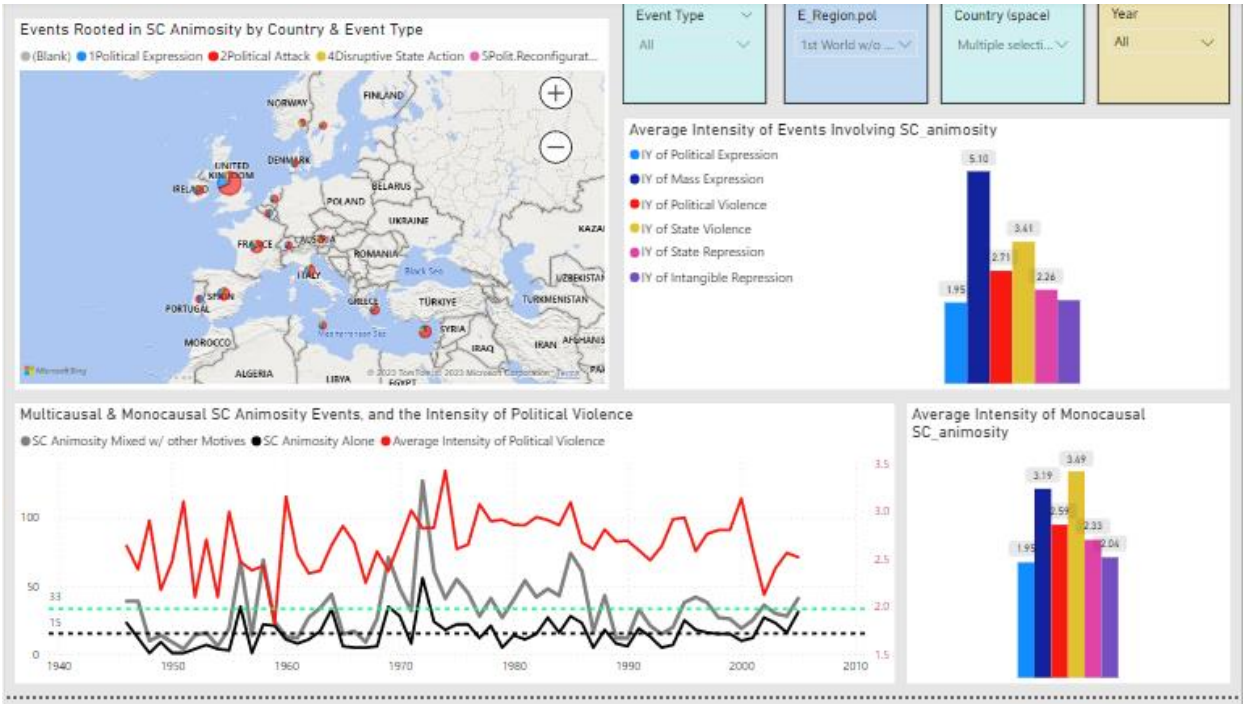
addressed and fixed, the antidote to communal violence is to institutionalize some regulated ways of non-violent conflict resolution.

The SPEED data, which spans 60 years, evidences some historical trends of the conflict events, and we know that societies thus far have always been segmented, with advantaged and disadvantaged groups of people pitted against each other. For millennia, wealth and legal servitude were the main group delimiters, followed by the importance of religion. The salience of other communal traits – language, ethnicity, region – increased when nation-states with unitary markets and educational systems started to form, and in this sense, they are not more "archaic" and "primitive" than capitalism itself. Even race's salience was brought about by colonialism and the American slavery. That is, most communal groups became self-aware of their group-ness in societies that promoted the enlightened/liberal ideology that all human beings are born free and equal. It should not come as a surprise that they have become vociferous about their sharing into economic, social, political, and cultural equality.

Looking at the divergent trajectories of the class conflicts and communal conflicts revealed by the SPEED data – class conflict events trending downward, sociocultural animosity events maintaining and even trending upward by 2005 – we could say that developed societies have been successful in institutionalizing ways in which the economically challenged may make their claims for more equality, yet they failed to work out good recipes for managing communal diversity. Actually, the European Union, and some other European countries, have reached solutions that keep away the old historical minorities from initiating newsworthy events, and, as opinion surveys attest, make them contented citizens of their country. The old continent's big problems are the newer immigrants and future immigration. These often elicit newsworthy activity fueled by anti-immigration animosity. The same anti-immigrant animosity has been present on other continents, as well, in the developed countries targeted by economic and political refugees, even in countries where the current dominant majority arrived as a very diverse immigrant population not that long ago. To this issue, there is no solution in sight, thus events rooted in this type of sociocultural animosity will probably persist in the near future, and mainly in the developed countries.

Yet in some respects the trajectory of sociocultural animosities resembles that of the class conflict. With the old historical minorities Europe achieved forms of co-existence where identities are decently accommodated and tensions are solved in institutionalized non-violent ways. And of course, more isolated examples of multi-ethnic/multi-national countries have long existed, such as Switzerland, Canada, and India. We would reckon that communal peace, as declining number and intensity of events rooted in sociocultural animosities, is possible, and does not need suppression of communal identities. Yet this claim takes us beyond what the SPEED data can directly support, since it does not distinguish between sociocultural animosities among long co-existing communal groups and animosity against non-citizens or new immigrant groups. Without this distinction, the occurrence and intensity of the communal conflict events in the European countries belonging to the "1st World" group shows some decline over the last two decades, but not a really convincing declining pattern.

Chart 5: Events rooted in sociocultural animosities in the European countries belonging to the "1st World" group



What the SPEED data can directly support, are the claims about the comparative destructiveness of events of different nature, and for the period between 1946-2005, events rooted in sociocultural animosities were not more frequent, more intense, or more escalating, than events rooted in other issues, such as animosity against government, class conflicts, or the strongmen's attempts to illegitimately retain power. Communal animosity events, on their own, were shown to be responsible for only 14% of all conflict events. SPEED also evidenced the multicausal origin of most newsworthy conflict events, and showed that communal conflicts are often associated with the groups' desire to obtain political rights.

REFERENCES

Cederman, Lars-Erik; Nils Weidmann & Kristian Skrede Gleditsch. 2011. Horizontal Inequalities and Ethno-nationalist Civil War: A Global Comparison, *American Political Science Review*, 105(3): 478–495.

Fearon, James & David D. Laitin. 1996. Explaining Interethnic Cooperation. *American Political Science Review*, 90(4): 715-35.

Flesken, Anaïd. 2018. Ethnic Parties, Ethnic Tensions? Results of an Original Election Panel Study. American Journal of Political Science, 62, 4, (967-981).

Gisselquist, Rachel M., Stefan Leiderer & Miguel Nino-Zarazua. 2016. Ethnic Heterogeneity and Public Goods Provision in Zambia: Evidence of a Subnational ''Diversity Dividend". *World Development*, Vol. 78, pp. 308–323, doi:10.1016/j.worlddev.2015.10.018.

Hillesund, Solveig; Karim Bahgat; Gray Barrett; Kendra Dupuy; Scott Gates; Håvard Mokleiv Nygård; Siri Aas Rustad; Håvard Strand; Henrik Urdal; & Gudrun Østby. 2018: Horizontal inequality and armed conflict: a comprehensive literature review. *Canadian Journal of Development Studies*, Vol. 39, Issue 4, pp. 463-480, https://doi.org/10.1080/02255189.2018.1517641.

Horowitz, Donald. 1985. *Ethnic Groups in Conflict: Theories, Patterns, and Policies*, University of California Press.

Ishiyama, John. 2009. Do Ethnic Parties Promote Minority Ethnic Conflict? *Nationalism & Ethnic Politics*. Vol. 15 Issue 1, p56-83.

Montalvo, José G., & Marta Reynal-Querol. 2021. Ethnic Diversity and Growth: Revisiting the Evidence. The Review of Economics and Statistics, 103 (3): 521–532.

Stewart, Frances (ed). 2008. *Horizontal inequalities and conflict: Understanding group violence in multiethnic societies*. Basingstoke [England]; New York: Palgrave Macmillan.

DATASETS

SPEED Global Random Sample 1945-2005 ("spp_public.xls"), Cline Center for Advanced Social Research, University of Illinois, data produced within the frames of the Social, Political, and Economic Event Database (SPEED) Project, https://clinecenter.illinois.edu/project/human-loop-event-data-projects/SPEED.

Historical Index of Ethnic Fractionalization (HIEF) data. The fractionalization indexes were calculated by Lenka Drazanova (https://openhumanitiesdata.metajnl.com/articles/10.5334/johd.16), based on the ethnic distribution data compiled by the Composition of Religious and Ethnic Groups (CREG) Project, also of the Cline Center for Advanced Social Research, https://clinecenter.illinois.edu/project/Religious-Ethnic-Identity/composition-religious-and-ethnic-groups-creg-project.

Teorell, Jan, Aksel Sundström, Sören Holmberg, Bo Rothstein, Natalia Alvarado Pachon, Cem Mert Dalli & Yente Meijers. 2023. *The Quality of Government Standard Dataset*, version Jan23. University of Gothenburg: The Quality of Government Institute, https://www.gu.se/en/quality-government doi:10.18157/qogstdjan23. We used the time series version.

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Jan Teorell, David Altman, Michael Bernhard, Agnes Cornell, M. Steven Fish, Lisa Gastaldi, Haakon Gjerløw, Adam Glynn, Ana Good God, Sandra Grahn, Allen Hicken, Katrin Kinzelbach, Joshua Krusell, Kyle L. Marquardt, Kelly McMann, Valeriya Mechkova, Juraj Medzihorsky, Natalia Natsika, Anja Neundorf, Pamela Paxton, Daniel Pemstein, Josefine Pernes, Oskar Rydén, Johannes von Römer, Brigitte Seim, Rachel Sigman, Svend-Erik Skaaning, Jeffrey Staton, Aksel Sundström, Eitan Tzelgov, Yi-ting Wang, Tore Wig, Steven Wilson and Daniel Ziblatt. 2023. "*V-Dem [Country-Year] Dataset v13*" Varieties of Democracy (V-Dem) Project. https://doi.org/10.23696/vdemds23.