Supplementary material

Additional information on the CPR data

The CPR data for Perm were collected by Andrei Semenov and colleagues through a commercial service that archives full versions of texts produced by newspapers (print and online), news agencies, as well as radio and TV stations in the post-Soviet space on both the federal and the regional level.[1] For Russia alone there are over 40,000 sources (Semenov 2017). The search string used was "акци* протест*".

Coding of CPR Data

The CPR from both Perm and Tyumen come with quotes from the original sources on the claims of protesters. This information was used to code the events according to the scheme developed by Lankina and colleagues and applied, e.g. in Lankina and Voznaya (2015). The authors suggest the following instructions for coding:

Category	*description*
Political	Political protest (1=yes; blank=otherwise) - Politically motivated anti-government and anti-regime protests at municipal, regional, and national levels. Anti-government protests with broad agendas (these protests may include other issues, but the protest against the regime/government itself or the protest for the protection of political rights form the crux of the protest); Protests dealing with electoral fraud; Protests calling for resignation of elected or politically appointed officials at all levels of government; Protests relating to political repressions and associated actions. These protests include protests for the release of those apprehended for political reasons (including protesters who were apprehended for protesting), memorial protests commemorating deaths related to political repression, protests in support of those who are undertaking political activism, protests against police abuse when dealing with repression of political activists; Protests against Russian foreign policy, such as protests against Russia's cooperation with Japan over the Kuril Islands, or protests in support of political events happening abroad, showing solidarity with protesters abroad.
Economic	Economic protest (1=yes; blank =otherwise). Protests against government economic policies, such as those affecting exchange rates, wages, etc.; wage- and worker-rights-related labour strikes
Social	Social protest (1=yes; blank =otherwise). Protests by, and specifically furthering the aims of, socially vulnerable groups of people such as pensioners, victims of Chernobyl, students, disabled people, people on state benefits
Legal	Cultural protest (1=yes; blank =otherwise). Protests against the destruction of monuments and of historically valuable buildings and sites; against change in city (area) names
Ecological	Legal protest (1=yes; blank =otherwise). Protests targeting unpopular legislation, its implementation (labour, criminal, and administrative codes); protest against illegal acts by state bodies or private companies (forced eviction, construction in inappropriate areas)
Cultural	Environmental protest (1=yes; blank =otherwise). Environmental issues, hazardous work conditions, waste dumping, destruction of forest reserves, parks, and protected woodlands

Notes: Coding scheme taken from Codebook available in Lankina (2018).

Additional figures

[1] Mustajoki, Arto. 2006. "The Integrum Database as a Powerful Tool in Research on Contemporary Russian." Integrum: *Tochnye metody i gumanitarnye nauki [Exact methods and the Humanities]*, edited by G. Nikiporets-Tagikava, pp. 50–76.

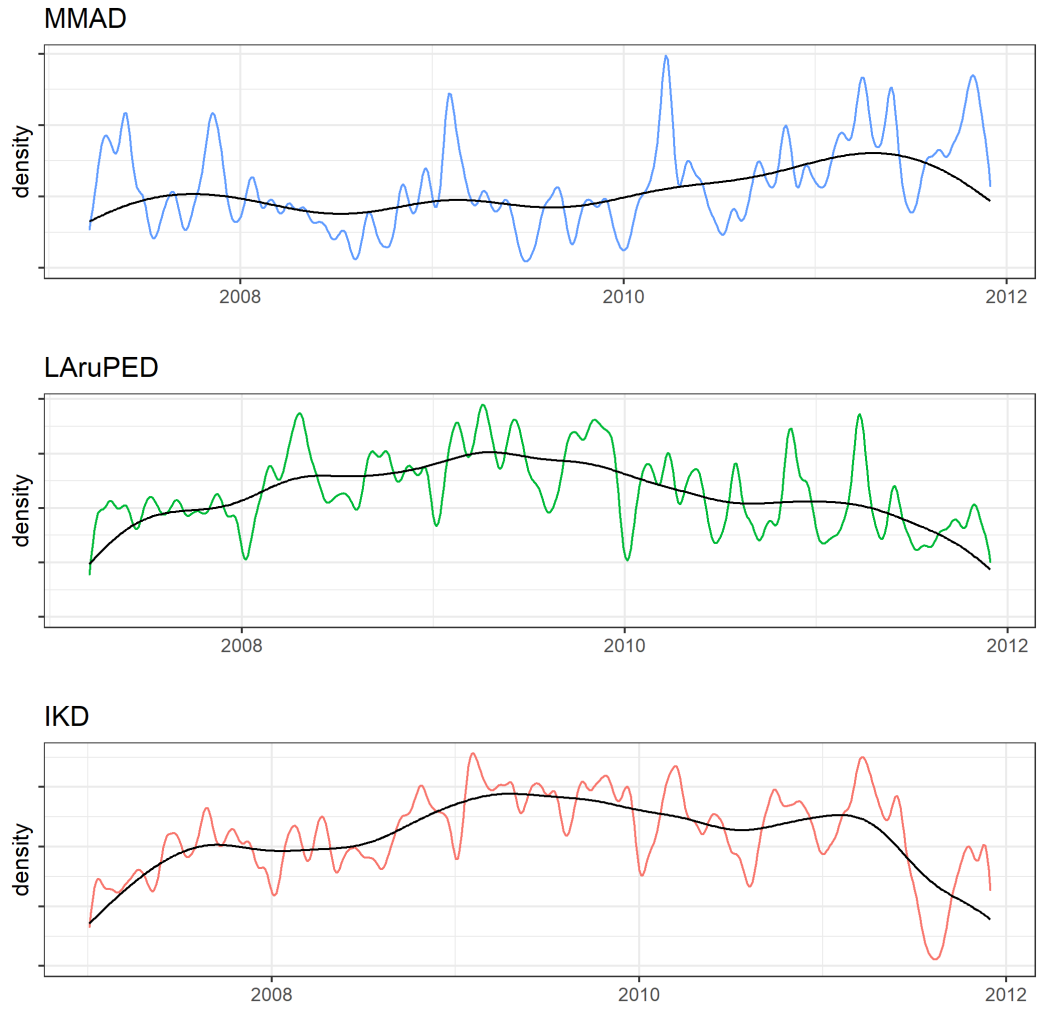

Figure A1. Protests events over time by data source, excluding period of "For Fair Elections" protest, with density curve and trend line (March 2007 – December 3, 2011).

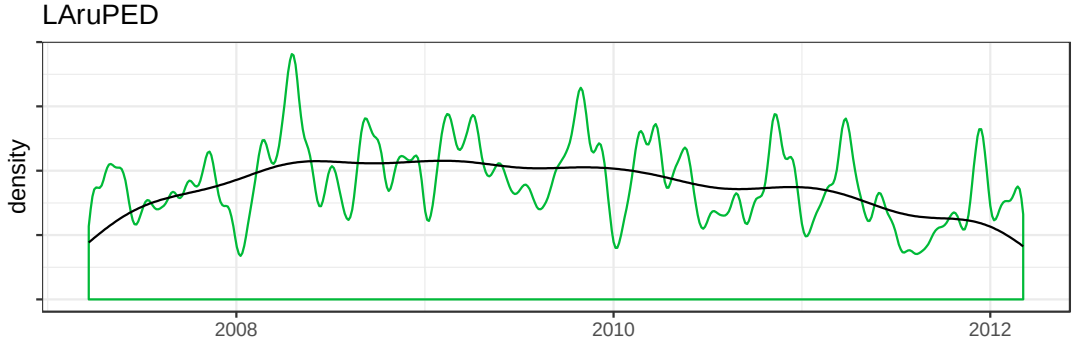

Figure A2. Protests events over time of LAruPED, subsetted to events with 25+ participants, with density curve and trend line (March 2007 – March 2012).

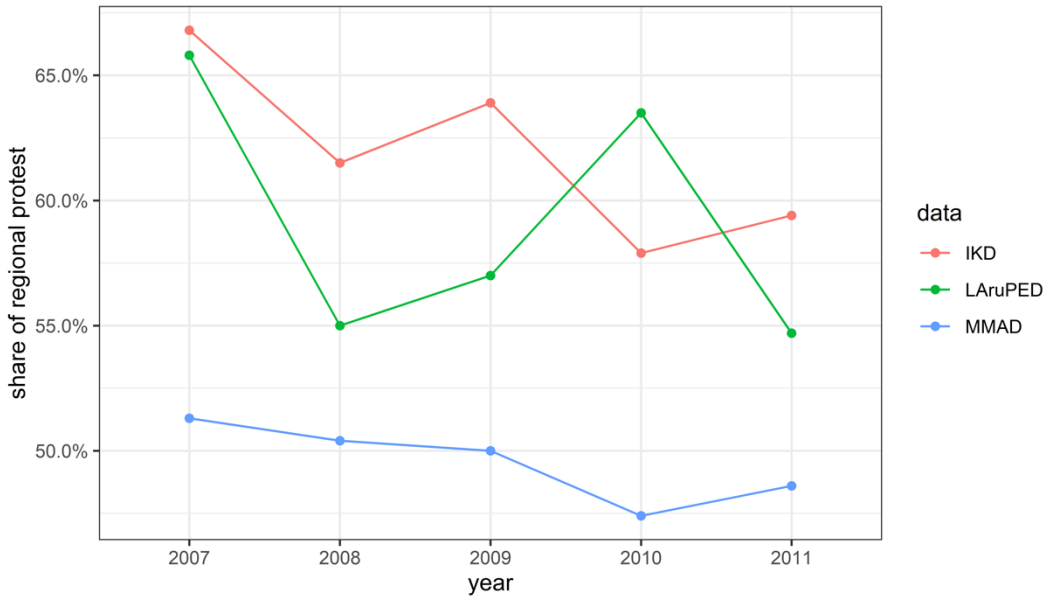

Figure A3. Share of protest events outside Moscow and St. Petersburg over time, by data source, FFE period excluded (March 2007 – December 3, 2011).

Additional tables

Table A1. Share of protest events above and below 25 participants in LAruPED, before and during For Fair Elections protest wave

	Before FFE	During FFE
Under 25 participants	23.9%	9.3%
25 and more participants	76.1%	90.7%

Notes: FFE = "For Fair Elections". Before FFE: March 2007 – December 3, 2011; during FFE: December 4, 2011 – March 5, 2012.

Treatment of duplicates and robustness check with reduced IKD data

A problem in PEA research is the existence of duplicates – events that are, by mistake, included more than once and thereby distort the overall picture. In this analysis, duplicates are dealt with in the following way.

First, all four data sets were checked for potential duplicates by listing the events that have the exact same values on different variables, like date, region, number of participants, and – if provided – protest topic (since the data sets no not include identical variables, this process diverged from data set to data set). In the case of MMAD, no potential duplicates were identified for the given period; in the case of LAruPED and CPR, all potential duplicates were checked manually. For CPR, the information on claims, organizers, and participants, that is given in the data set, was used to identify actual duplicates, which led to the exclusion of 4 events (out of 462). For LAruPED, the manual check used the URLs to the original sources, leading to the exclusion of 12 events (out of 4509). All calculations and figures in the manuscript and in this appendix are based on these cleaned data sets.

For the IKD data, the initial procedure revealed 166 potential duplicates (out of 5593 events in the analyzed period). In contrast to the other data sets, however, there was no possibility to manually check whether these were indeed duplicates, because the data do not provide further information on the events. Given that the manual checks for LAruPED revealed a very low number of actual duplicates, the analysis in the manuscript proceeds without further changes to the IKD data, assuming that no duplicates exist. However, below I provide a robustness check for all calculations and figures with the 166 potential duplicates excluded (the calculation of overlap is not reproduced as it yields identical results). The results in no way change the conclusions drawn in the study.

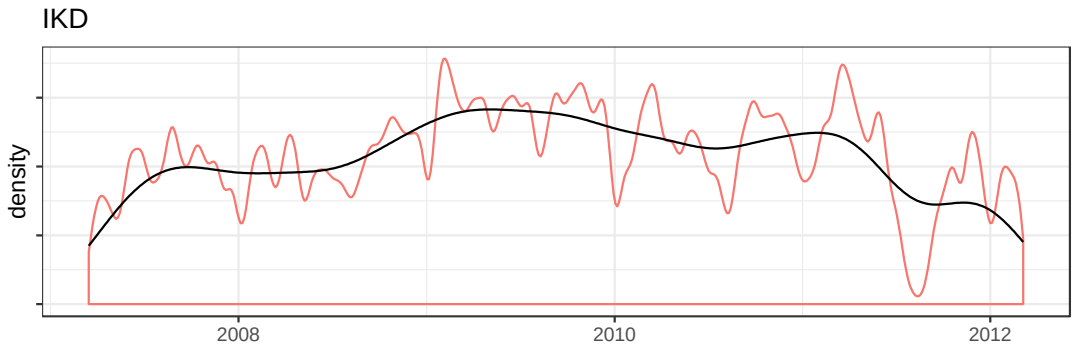

Figure A4. Protests events over time of IKD data, 166 potential duplicates dropped, with density curve and trend line, (March 2007 – March 2012).

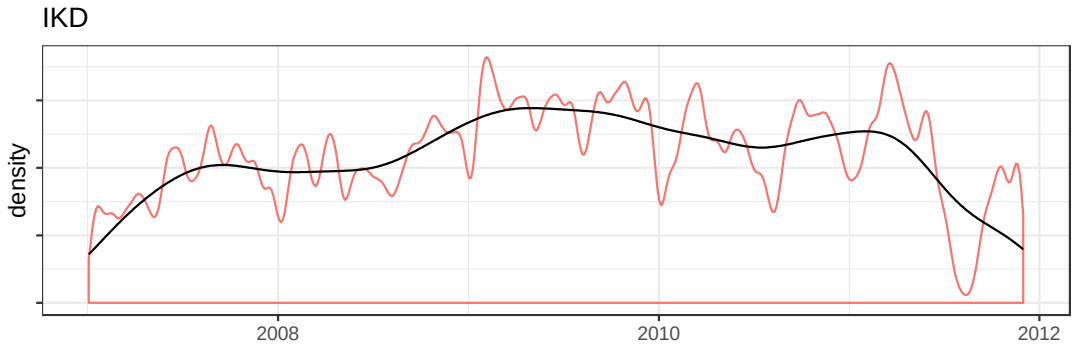

Figure A5. Protests events over time of IKD data, 166 potential duplicates dropped, excluding period of "For Fair Elections" protest, with density curve and trend line (March 2007 – December 3, 2011).

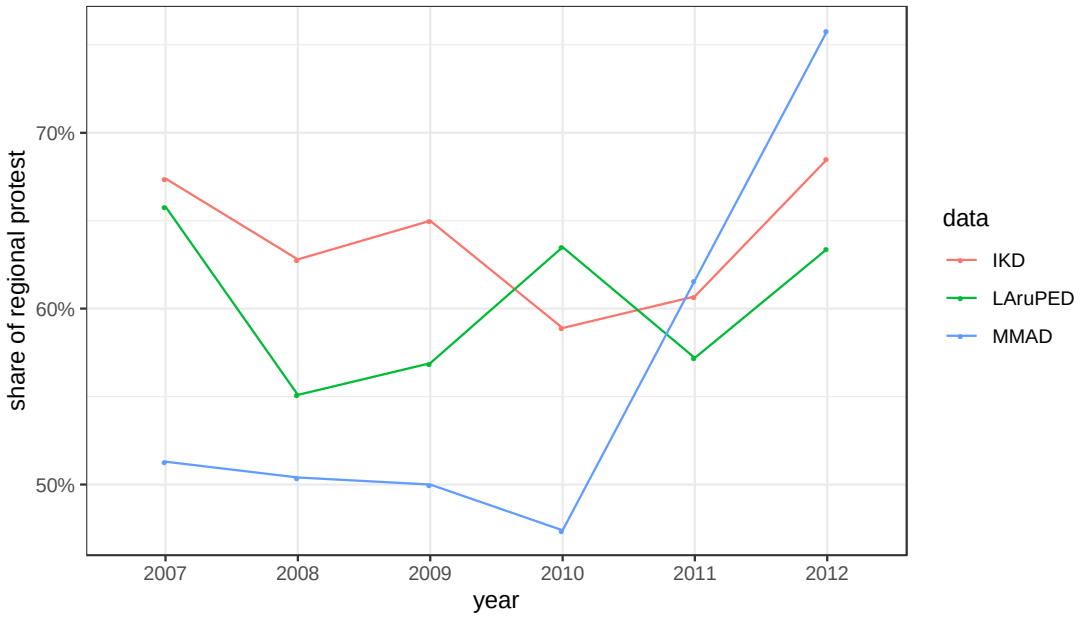

Figure A6. Share of protest events outside Moscow and St. Petersburg over time, by data source (March 2007 – March 2012). 166 potential duplicates In the IKD data were dropped.

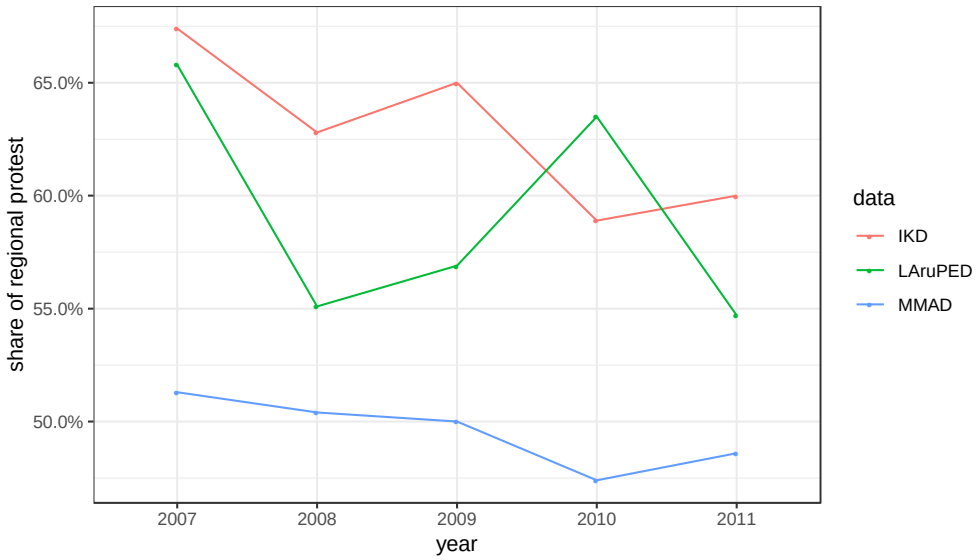

Figure A7. Share of protest events outside Moscow and St. Petersburg over time, by data source, FFE period excluded (March 2007 – December 3, 2011). 166 potential duplicates in the IKD data were dropped.

Table A2. Regions with the highest and the lowest number of events per data source (March 16, 2007 – March 5, 2012), with 166 potential duplicates in the IKD data excluded.

MMAD			**LAruPED**			**IKD**		
Region	*# of events*	*% of total*	*Region*	*# of events*	*% of total*	*Region*	*# of events*	*% of total*
Moscow City	346	30.0%	Moscow City	1288	29.6%	Moscow City	1680	31.4%
St. Petersburg	152	13.2%	St. Petersburg	484	11.1%	St. Petersburg	301	5.6%
Dagestan	54	4.7%	Samara	176	4.0%	Leningrad	225	4.2%
Primorie	40	3.5%	Moscow Oblast	130	3.0%	Novosibirsk	163	3.0%
Sverdlovsk	37	3.2%	Penza	115	2.6%	Sverdlovsk	139	2.6%
Novosibirsk	27	2.3%	Sverdlovsk	104	2.4%	Samara	136	2.5%
Kaliningrad	23	2.0%	Kaliningrad	99	2.3%	Irkutsk	130	2.4%
Samara	20	1.7%	Voronezh	99	2.3%	Moscow Oblast	126	2.4%
Bashkortostan	19	1.6%	Primorie	96	2.2%	Kaliningrad	124	2.3%
North Ossetia	19	1.6%	Kirov	89	2.0%	Chelyabinsk	121	2.3%

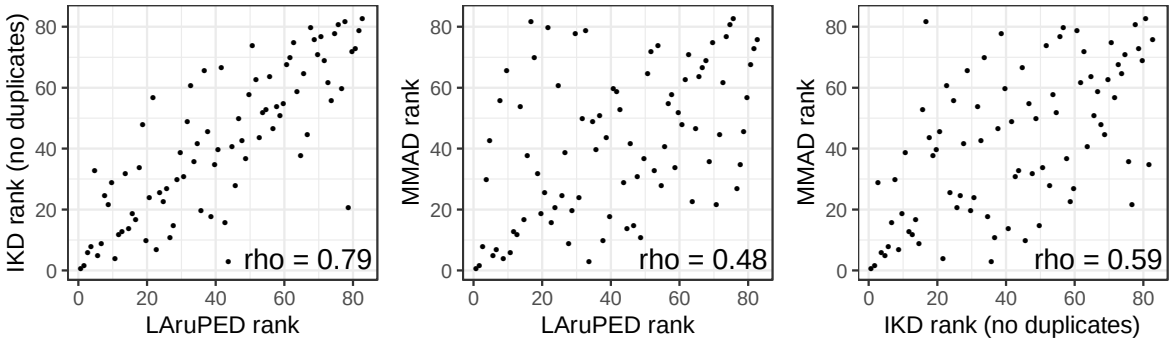

Figure A8. Rank correlations of the number of covered protest events per region, by combination of data sets. Period covered: March 2007 – March 2012, with 166 potential duplicates in the IKD data excluded.

Table A3. Protest events in North Caucasus by data source and year, with 166 potential duplicates in the IKD data excluded.

	2007	**2008**	**2009**	**2010**	**2011**	**2012**
MMAD	25 (31.6%)	31 (51.7%)	10 (14.3%)	11 (11.8%)	32 (12.4%)	5 (5.3%)
LAruPED	18 (4.6%)	16 (3.0%)	17 (2.7%)	21 (4.0%)	8 (2.0%)	3 (3.5%)
IKD	14 (2.8%)	19 (3.0%)	15 (1.7%)	17 (2.6%)	4 (0.7%)	1 (1.0%)

Notes: Cells show absolute numbers of reported events in North Caucasus and the share among all regional protest (in brackets). North Caucasus defined as Chechnya, Dagestan, Ingushetia, Kabardino-Balkaria, Karachaevo-Cherkessia, and North Ossetia.

Comparison of covered topics in LAruPED and IKD

The following two figures directly compare the thematic coverage of LAruPEd and IKD. For that, the IKD topics – which are given in 26 categories – were recoded according to the six thematic categories used by Lankina (see Lankina and Voznaya 2015). This likely includes some information loss, including the problem of dealing with nationalist/ethnic protest, so that about 2.5% of the protest events in IKD could not be assigned a category. However, the comparison still supports the expectation that the two data sets primarily differ on their respective share of political and social protest – with the other categories showing very similar shares across the data sets.

The IKD categories were recoded as follows:

IKD category ("What action would it take to satisfy demands?"	**recoded to**
Personnel changes	political
Elections irregularities	political
National festival	cultural
Improved wages/ working conditions	economic
Commercial/market related demands	economic
Change in ownership/control of enterprise	economic
Labor rights	economic
Environmental	environmental
Enforcement of the law: physical security	legal
Criminal justice	legal
Freedom of Association/Civil Rights	legal
Corruption/police issues	legal
Gay rights	legal
Women's rights	legal
Enforcement of the law: material	social
Increased social spending/change in material distribution	social
Construction/Development Issues/Housing	social
Healthcare	social
Historical commemoration	cultural
Education	cultural
Ethnic politics (removal, resettlement, land claim, exclusion of particular ethnic groups)	NA
Foreign affairs	NA
Nationalist demands	NA
Other	NA
Military reform	NA
Caucasus war	NA
Prisons	NA
Drivers' Rights	NA

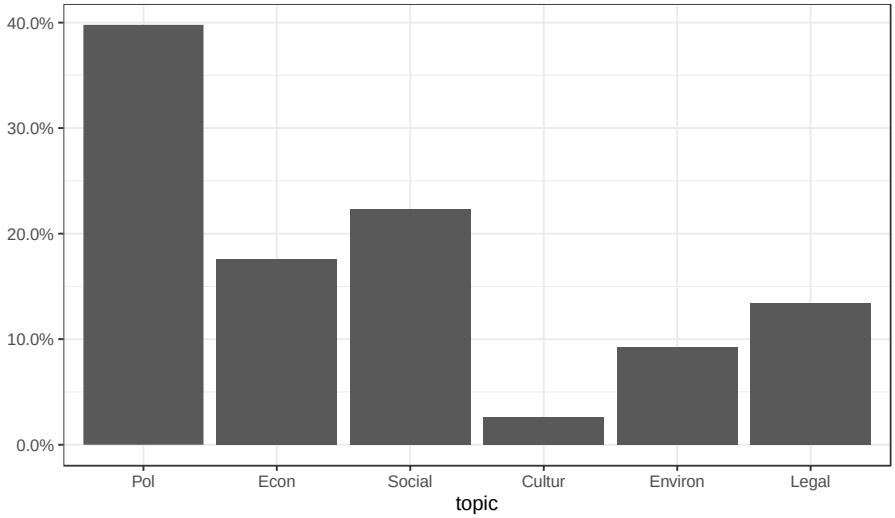

Figure A9. Share of protest events by thematic category in LAruPED.

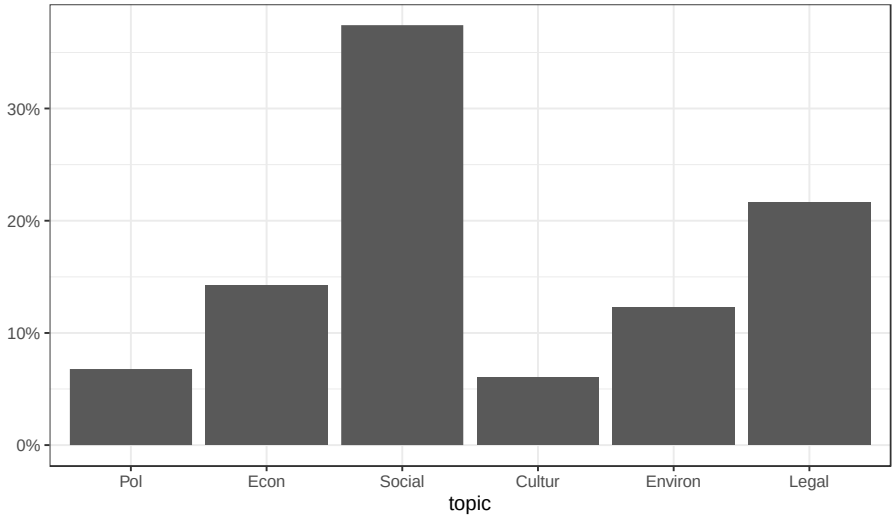

Figure A10. Share of protest events by thematic category in IKD. Categories recoded according to scheme displayed above.

Manual check of overlap between LAruPEd and CPR in Perm and Tyumen

Since the IKD data does not provide access to the original sources or any other substantive information on the coded events, it is not possible to manually check whether the events coded as overlapping with other data sets are in fact the same events. However, since LAruPED does provide access to original sources and CPR includes detailed information on each protest event directly drawn from the underlying sources, a comparison of factual overlap between these two data sets in Perm and Tyumen is possible. The check revealed that in about three quarters of cases (23 of 30 events) that were identified as overlapping by the method described in the manuscript (see also figure 4 in the manuscript), the original sources confirmed that the two were indeed the same event. On the one hand, this result shows that the crude method used – counting as same events when date and region are identical – produces, as expected, a substantial number of false positives. On the other hand, the numbers do show that the method has a sound empirical base and can serve as a first approximation – in particular when the method is used consciously as a most-likely case to detect overlap.