SUPPLEMENTARY MATERIALS - PART I

Supplementary materials to: Lenka Bustikova, David Siroky, Saud Alashri, Sultan Alzahrani. Predicting Partisan Responsiveness: A Probabilistic Text Mining Time-Series Approach. *Political Analysis*.

Replication

In this appendix we describe the general procedure for replicating the analysis described in the paper. We first describe the required tools, and then the sequence of steps from storing the data into database to viewing the results. The source code and discussion assumes some experience with the following tools and programming languages: *PostgreSQL, Apache SOLR, Java (Eclipse recommended), Matlab, R.* All source code is accessible online [anonymized]:

Detailed description of how to run the each step is given in the README files.

1. Creating Database to store documents (PostgreSQL).
2. Indexing the data from PostgreSQL into Apache SOLR for fast retrieval and processing.
3. Running LDA to generate topics.
4. Determining relatedness of spikes using Threshold determination script.
5. Generating frames for each camp based on each issue.
6. Using frames to train logistic regression classifier and predict escalation.

========== >>(ReadMe)<< ==========

Predicting Partisan Responsiveness READ ME Replication

==========

This project aims to determine contentious frames between opposing camps in Slovakia. Then, use them to predict online escalation. Please refer to the paper for further details about the methodology.

==========

Getting Started:
....These instructions will get you a copy of the project up and running on your local machine for
....development and testing purposes.

==========

Prerequisites:
You need the following tools/languages:
- PostgreSQL.
- Apache SOLR.
- Java (Eclipse).
- Matlab.
- R.

==========

Experiment steps:

1- Database:
....Please create database tables using the following backup file: slovakiaDB. You can also create
....the tables and store the data by these two sql scripts:

1.1 Slovakia Table:

```
CREATE TABLE Slovakia
(
....id integer NOT NULL,
....file_name text,
....file_content text,
....article_date date,
CONSTRAINT "Slovakia_pkey" PRIMARY KEY (id)
)
WITH (
....OIDS=FALSE
);
ALTER TABLE Slovakia
OWNER TO postgres;
```

1.2 Slovakia parties:

```
CREATE TABLE Slovakia_Org
(
....tid integer NOT NULL,
....org_name text,
....org_type text,
....country text,
....CONSTRAINT "Slovakia_Org_pkey" PRIMARY KEY (tid),
....CONSTRAINT "Slovakia_Org_tid_fkey" FOREIGN KEY (tid)
....REFERENCES Slovakia (id) MATCH SIMPLE
....ON UPDATE NO ACTION ON DELETE NO ACTION
)
WITH (
....OIDS=FALSE
);
ALTER TABLE Slovakia_Org
OWNER TO postgres;
```

==========

2- Data:
....2.1 Data files are located under folder (Data).

==========

3- Apache SOLR:
....3.1 Download Solr and extract it to a directory of your choice,
........(for installation guidance, you may follow to this example:
........http://lucene.apache.org/solr/quickstart.html).
....3.2 Once the server installed, configure SOLR by adjusting these files based on your database:
....3.2.1 under ../example/solr/collection1/conf/ modify the following:
........A. add your database connection info to solrconfig.xml
........B. copy schema.xml (from solr_config folder on dropbox) into this folder and change the
..........required directories.
........C. copy data-config.xml (from solr_config folder on dropbox) into this folder.
....3.3 Next, index the data from postgreSQL to SOLR.

==========

4- Mallet LDA:
....4.1 Install LDA from (http://mallet.cs.umass.edu/).
....4.2 For each grid/group issue, run LDA (under LDA experiment folder) on the entire corpora
........to generate 100 topics each with 20 terms and save them with their distributions.

==========

```
5- Utilize the Threshold determination R script under folder (Thresholds_determination) to get
....the thresholds for spikes.

=============================================================================================

6- Train and test the classifier using logistic regression:
....5.1 Download the SLEP package from:  http://www.yelab.net/software/SLEP/
....5.2 Replace the logistic.m file with the one located under (Frame_Identification_and_classification)
........folder.
....5.3 Adjust the directories as needed.

=============================================================================================
* Copyright (C) 2017 - Code released under the MIT License.
```

Replicating this pipeline and generating accurate spike pair categorization using the same or different official websites is straightforward if one follows the following already commonplace practices:

- Remove stopword, punctuation, websites domain names, meaningless frequent keywords, etc., as initial text preparation.
- When converting unstructured articles of size n into vector space model (VSM) of size $(n \times m)$, m has to be confined by n such that $(m << n)$. We maintain m to n as $m : n = 1 : 4$, and our experiments yield m and n as 10,597 and 2,500 respectively.
- LDA requires multiple runs, and adopting model selection criteria, such as Elbow methods, can help in deducing initial k settings. Then, interim LDA results with various: $k_{-i}, \ldots, k, \ldots, k_{+i}$ topic dimensions can be evaluated qualitatively for the final k selection - all model selection approaches can be adapted to figure out initial settings, but cannot replace the manual qualitative validation phase for the final k selection. Experimentally, results of larger k values result in less mutated results when other settings–e.g., priors parameters and the maximum number of iterations–are kept fixed.
- Keep default value of logistic regression as it is employing a gradient descent search, namely max number of iteration of **1k**, use the smallest possible stopping criteria tolerance to achieve the highest approximation of the classifier hyper parameters.

Software License

Grid-Group Definition of Political Parties

Subject matter experts were guided by the grid-group dimensions of social control theory (Douglas and Wildavsky 1982, Douglas 1970) in selecting political parties to be included and in identifying *key issues* in their corpus. The *grid-group* typology characterizes four ideal type modes of social action along a grid-group axis. The grid dimension captures policy positions of radical actors that are compatible with authoritarianism and social and cultural conservatism. In its pure form, the grid dimension has no ethnic basis. An actor that promotes law and order, along with uncritical obedience to authority, religious or secular, is classified as high on the grid dimension.In policy terms, the high grid dimension is associated with support for a high degree of state intrusion into the daily life of its citizens and a 'strongman' position in international affairs. The group dimension captures nationalism and is therefore associated with exclusionary ethnicity-based appeals. It conceptualizes identity in terms of 'the ethnic other' and is grounded in a distinction between the in-group and the out-group. An actor that propagates nationalism on behalf of the titular nationality would score high on the group dimension. Core-group ethnocentric, anti-minorities and socially conservative actors are classified as radicals.

Source:

Bustikova, L. 2018. Radical Right Mobilization in Eastern Europe. Cambridge University Press (under contract).

Douglas, M. (1970). Natural Symbols. Explorations in Cosmology. London: Routledge.

Douglas, M. and Wildavsky, A. (1982). Risk and Culture: An Essay on the Selection of Technical and Environmental Dangers. Berkeley, CA: University of California Press.

Predictions of Spikes

Table 1: Predictions for *escalated* ethnic spikes

Issue	Dimension	TP	FP	Precision	Recall	F-measure
Minorities	Group	0.79	0.15	0.82	0.79	0.81
Nation	Group	0.95	0.29	0.82	0.95	0.88
Language	Group	0.88	0.19	0.83	0.88	0.86
EU/Enlargement	Grid	0.77	0.12	0.87	0.77	0.82
Economics	Grid	0.86	0.26	0.78	0.86	0.82
Interstate	Grid	0.92	0.25	0.86	0.92	0.89

Table 2: Predictions for *ignored* ethnic spikes

Issue	Dimension	TP	FP	Precision	Recall	F-measure
Minorities	Group	0.85	0.21	0.82	0.85	0.84
Nation	Group	0.71	0.042	0.92	0.71	0.80
Language	Group	0.81	0.12	0.86	0.81	0.83
EU/Enlargement	Grid	0.87	0.23	0.78	0.87	0.82
Economics	Grid	0.74	0.14	0.82	0.74	0.78
Interstate	Grid	0.75	0.08	0.85	0.75	0.80

Table 3: Predictions for *escalated* radical spikes

Issue	Dimension	TP	FP	Precision	Recall	F-measure
Minorities	Group	0.83	0.13	0.85	0.83	0.84
Nation	Group	0.79	0.19	0.81	0.79	0.80
Language	Group	0.80	0.16	0.84	0.80	0.82
EU/Enlargement	Grid	0.84	0.22	0.83	0.84	0.84
Economics	Grid	0.82	0.15	0.86	0.82	0.84
Interstate	Grid	0.85	0.15	0.87	0.85	0.86

Table 4: Predictions for *ignored* radical spikes

Issue	Dimension	TP	FP	Precision	Recall	F-measure
Minorities	Group	0.87	0.17	0.85	0.87	0.86
Nation	Group	0.81	0.20	0.79	0.81	0.80
Language	Group	0.83	0.20	0.80	0.83	0.82
EU/Enlargement	Grid	0.78	0.16	0.79	0.78	0.78
Economics	Grid	0.85	0.18	0.80	0.85	0.82
Interstate	Grid	0.85	0.15	0.83	0.85	0.84

Naive Bayes and Random Forest Classifiers

We experimented with various classifiers and reported the best performing ones: Naive Bayes and Random Forest. SLEP outperformed both Naive Bayes and Random Forest classifiers (F-measure). Generally, Random Forest outperformed Naive Bayes, and both classifiers performed better in predicting escalated spikes compared to ignored spikes.

Classifier: Naive Bayes

Escalated Ethnic

Issue	Dimension	TP	FP	Precision	Recall	F-measure
Minorities	Group	0.753	0.238	0.762	0.753	0.757
Nation	Group	0.841	0.250	0.779	0.841	0.809
Language	Group	0.777	0.240	0.791	0.777	0.784
EU/Enlargement	Grid	0.786	0.182	0.846	0.786	0.815
Economics	Grid	0.810	0.275	0.756	0.810	0.782
Interstate	Grid	0.889	0.222	0.800	0.889	0.842

Ignored Ethnic

Issue	Dimension	TP	FP	Precision	Recall	F-measure
Minorities	Group	0.762	0.247	0.753	0.762	0.757
Nation	Group	0.750	0.159	0.818	0.750	0.783
Language	Group	0.760	0.223	0.745	0.760	0.753
EU/Enlargement	Grid	0.818	0.214	0.750	0.818	0.783
Economics	Grid	0.725	0.190	0.784	0.725	0.753
Interstate	Grid	0.778	0.111	0.875	0.778	0.824

Escalated Radical

Issue	Dimension	TP	FP	Precision	Recall	F-measure
Minorities	Group	0.790	0.298	0.721	0.790	0.754
Nation	Group	0.739	0.373	0.676	0.739	0.706
Language	Group	0.657	0.204	0.770	0.657	0.709
EU/Enlargement	Grid	0.713	0.351	0.694	0.713	0.703
Economics	Grid	0.817	0.214	0.831	0.817	0.824
Interstate	Grid	0.897	0.380	0.763	0.897	0.825

Ignored Radical

Issue	Dimension	TP	FP	Precision	Recall	F-measure
Minorities	Group	0.702	0.210	0.775	0.702	0.737
Nation	Group	0.627	0.261	0.695	0.627	0.659
Language	Group	0.796	0.343	0.690	0.796	0.739
EU/Enlargement	Grid	0.649	0.287	0.670	0.649	0.660
Economics	Grid	0.786	0.183	0.769	0.786	0.777
Interstate	Grid	0.620	0.103	0.815	0.620	0.704

Figure 1: Naive Bayes Classifier

Classifier: Random Forest

Escalated Ethnic

Issue	Dimension	TP	FP	Precision	Recall	F-measure
Minorities	Group	0.824	0.429	0.660	0.824	0.733
Nation	Group	0.857	0.250	0.783	0.857	0.818
Language	Group	0.875	0.292	0.778	0.875	0.824
EU/Enlargement	Grid	0.893	0.591	0.658	0.893	0.758
Economics	Grid	0.857	0.250	0.783	0.857	0.818
Interstate	Grid	0.800	0.083	0.889	0.800	0.842

Ignored Ethnic

Issue	Dimension	TP	FP	Precision	Recall	F-measure
Minorities	Group	0.571	0.176	0.762	0.571	0.653
Nation	Group	0.750	0.143	0.833	0.750	0.789
Language	Group	0.708	0.125	0.829	0.708	0.764
EU/Enlargement	Grid	0.409	0.107	0.750	0.409	0.529
Economics	Grid	0.750	0.143	0.833	0.750	0.789
Interstate	Grid	0.917	0.200	0.846	0.917	0.880

Escalated Radical

Issue	Dimension	TP	FP	Precision	Recall	F-measure
Minorities	Group	0.645	0.220	0.741	0.645	0.690
Nation	Group	0.783	0.446	0.649	0.783	0.710
Language	Group	0.725	0.127	0.805	0.725	0.763
EU/Enlargement	Grid	0.880	0.320	0.754	0.880	0.812
Economics	Grid	0.874	0.333	0.720	0.874	0.789
Interstate	Grid	0.876	0.254	0.825	0.876	0.850

Ignored Radical

Issue	Dimension	TP	FP	Precision	Recall	F-measure
Minorities	Group	0.780	0.355	0.693	0.780	0.734
Nation	Group	0.554	0.217	0.709	0.554	0.622
Language	Group	0.873	0.275	0.815	0.873	0.843
EU/Enlargement	Grid	0.680	0.120	0.835	0.680	0.750
Economics	Grid	0.667	0.126	0.843	0.667	0.745
Interstate	Grid	0.746	0.124	0.815	0.746	0.779

Figure 2: Random Forest Classifier

LDA comparison to Vector Space and Word Embedding

Escalated Ethnic

Issues	Dimension	TP	FP	Precision	Recall	F-measure
Minorities	Group	0.79	0.15	0.82	0.79	0.81
Nation	Group	0.95	0.29	0.82	0.95	0.88
Language	Group	0.88	0.19	0.83	0.88	0.86
EU/Enlargement	Grid	0.77	0.12	0.87	0.77	0.82
Economic	Grid	0.86	0.26	78	0.86	0.82
Interstate	Grid	0.92	0.25	0.86	0.92	0.89

Ignored Ethnic

Issues	Dimension	TP	FP	Precision	Recall	F-measure
Minorities	Group	0.85	0.21	0.82	0.85	0.84
Nation	Group	0.71	0.05	0.92	0.71	0.8
Language	Group	0.81	0.12	0.86	0.81	0.83
EU/Enlargement	Grid	0.87	0.23	0.78	0.87	0.82
Economic	Grid	0.74	0.14	0.82	0.74	0.78
Interstate	Grid	0.75	0.08	0.85	0.75	0.8

Escalated Radical

Issues	Dimension	TP	FP	Precision	Recall	F-measure
Minorities	Group	0.83	0.13	0.85	0.83	0.84
Nation	Group	0.79	0.19	0.81	0.79	0.8
Language	Group	0.8	0.16	0.84	0.8	0.82
EU/Enlargement	Grid	0.84	0.22	0.83	0.84	0.84
Economic	Grid	0.82	0.15	0.86	0.82	0.84
Interstate	Grid	0.85	0.15	0.87	0.85	0.86

Ignored Radical

Issues	Dimension	TP	FP	Precision	Recall	F-measure
Minorities	Group	0.87	0.17	0.85	0.87	0.86
Nation	Group	0.81	0.2	0.79	0.81	0.8
Language	Group	0.83	0.2	0.8	0.83	0.82
EU/Enlargement	Grid	0.78	0.16	0.79	0.78	0.78
Economic	Grid	0.85	0.18	0.8	0.85	0.82
Interstate	Grid	0.85	0.15	0.83	0.85	0.84

Figure 3: LDA

Escalated Ethnic

Issues	Dimension	TP	FP	Precision	Recall	F-measure
Minorities	Group	0.64	0.26	0.68	0.74	0.71
Nation	Group	0.69	0.13	0.74	0.87	0.80
Language	Group	0.63	0.33	0.65	0.67	0.66
EU/Enlargement	Grid	0.76	0.22	0.76	0.78	0.77
Economic	Grid	0.30	0.08	0.63	0.92	0.75
Interstate	Grid	0.66	0.31	0.67	0.69	0.68

Ignored Ethnic

Issues	Dimension	TP	FP	Precision	Recall	F-measure
Minorities	Group	0.74	0.36	0.71	0.64	0.67
Nation	Group	0.87	0.31	0.84	0.69	0.76
Language	Group	0.67	0.37	0.65	0.63	0.64
EU/Enlargement	Grid	0.78	0.24	0.77	0.76	0.76
Economic	Grid	0.92	0.70	0.75	0.30	0.43
Interstate	Grid	0.69	0.34	0.68	0.66	0.67

Escalated Radical

Issues	Dimension	TP	FP	Precision	Recall	F-measure
Minorities	Group	0.43	0.35	0.54	0.65	0.59
Nation	Group	0.57	0.37	0.60	0.63	0.61
Language	Group	0.55	0.42	0.57	0.58	0.58
EU/Enlargement	Grid	0.50	0.40	0.55	0.60	0.57
Economic	Grid	0.33	0.00	0.91	1.00	0.95
Interstate	Grid	0.50	0.36	0.64	0.64	0.64

Ignored Radical

Issues	Dimension	TP	FP	Precision	Recall	F-measure
Minorities	Group	0.65	0.57	0.55	0.43	0.48
Nation	Group	0.63	0.43	0.60	0.57	0.59
Language	Group	0.58	0.45	0.57	0.55	0.56
EU/Enlargement	Grid	0.60	0.50	0.55	0.50	0.52
Economic	Grid	1.00	0.67	1.00	0.33	0.50
Interstate	Grid	0.64	0.50	0.50	0.50	0.50

Figure 4: Vector Space Model

Escalated Ethnic

Issues	Dimension	TP	FP	Precision	Recall	F-measure
Minorities	Group	0.79	0.21	0.77	0.79	0.78
Nation	Group	0.68	0.21	0.77	0.68	0.72
Language	Group	0.77	0.21	0.77	0.77	0.77
EU/Enlargement	Grid	0.85	0.18	0.87	0.85	0.86
Economic	Grid	0.68	0.21	0.75	0.68	0.72
Interstate	Grid	0.82	0.23	0.82	0.82	0.82

Ignored Ethnic

Issues	Dimension	TP	FP	Precision	Recall	F-measure
Minorities	Group	0.79	0.21	0.81	0.79	0.80
Nation	Group	0.79	0.32	0.71	0.79	0.75
Language	Group	0.79	0.23	0.79	0.79	0.79
EU/Enlargement	Grid	0.82	0.15	0.80	0.82	0.81
Economic	Grid	0.79	0.32	0.72	0.79	0.75
Interstate	Grid	0.77	0.18	0.77	0.77	0.77

Escalated Radical

Issues	Dimension	TP	FP	Precision	Recall	F-measure
Minorities	Group	0.80	0.18	0.84	0.80	0.82
Nation	Group	0.78	0.18	0.83	0.78	0.80
Language	Group	0.79	0.21	0.78	0.79	0.79
EU/Enlargement	Grid	0.81	0.17	0.82	0.81	0.81
Economic	Grid	0.81	0.17	0.82	0.81	0.81
Interstate	Grid	0.79	0.17	0.81	0.79	0.80

Ignored Radical

Issues	Dimension	TP	FP	Precision	Recall	F-measure
Minorities	Group	0.82	0.20	0.77	0.82	0.79
Nation	Group	0.82	0.22	0.76	0.82	0.79
Language	Group	0.79	0.21	0.80	0.79	0.80
EU/Enlargement	Grid	0.83	0.19	0.82	0.83	0.82
Economic	Grid	0.83	0.19	0.82	0.83	0.82
Interstate	Grid	0.83	0.21	0.81	0.83	0.82

Figure 5: Word Embedding

Paired t-test comparisons

	LDA (F-measures)	*LEX Base model (F-measures)*
Mean	0.827916667	0.64113716
Variance	0.00081721	0.01443796
Observations	24	24
Pearson Correlation	0.23975323	
Hypothesized Mean Difference	0	
df	23	
t Stat	7.843975497	Reject Ho and Accept the Ha (alpha = 0.01)
P(T<=t) one-tail	3.00173E-08	
t Critical one-tail	1.713871528	
P(T<=t) two-tail	6.00346E-08	
t Critical two-tail	2.06865761	

Figure 6: Paired t-test comparison: LDA and Baseline Lexical Model (LEX)

	LDA (F-measures)	*Word Embedding (F-measures)*
Mean	0.827916667	0.792704568
Variance	0.00081721	0.001102268
Observations	24	24
Pearson Correlation	0.019064237	
Hypothesized Mean Difference	0	
df	23	
t Stat	3.975016046	Reject Ho and Accept the Ha (alpha = 0.01)
P(T<=t) one-tail	0.000299269	
t Critical one-tail	1.713871528	
P(T<=t) two-tail	0.000598538	
t Critical two-tail	2.06865761	

Figure 7: Paired t-test comparison: LDA and Word Embedding (EMB)

	Word Embedding (F-measures)	*LEX Base model (F-measures)*
Mean	0.792704568	0.64113716
Variance	0.001102268	0.01443796
Observations	24	24
Pearson Correlation	-0.157143363	
Hypothesized Mean Difference	0	
df	23	
t Stat	5.729730863	Reject Ho and Accept the Ha (alpha = 0.01)
P(T<=t) one-tail	3.89101E-06	
t Critical one-tail	1.713871528	
P(T<=t) two-tail	7.78202E-06	
t Critical two-tail	2.06865761	

Figure 8: Paired t-test comparison: Baseline Lexical Model (LEX) and Word Embedding (EMB)

	F-measures		
	LEX (F-measures) Base model	Word Embedding doc2vector (F-measures)	LDA (F-measures)
	0.71	0.78	0.81
	0.80	0.72	0.88
	0.66	0.77	0.86
	0.77	0.86	0.82
	0.75	0.72	0.82
	0.68	0.82	0.89
	0.67	0.80	0.84
	0.76	0.75	0.8
	0.64	0.79	0.83
	0.76	0.81	0.82
	0.43	0.75	0.78
	0.67	0.77	0.8
	0.59	0.82	0.84
	0.61	0.80	0.8
	0.58	0.79	0.82
	0.57	0.81	0.84
	0.95	0.81	0.84
	0.64	0.80	0.86
	0.48	0.79	0.86
	0.59	0.79	0.8
	0.56	0.80	0.82
	0.52	0.82	0.78
	0.50	0.82	0.82
	0.50	0.82	0.84
MEAN	0.64	0.79	0.83
STD	0.12	0.03	0.03

Figure 9: F-Measures

SUPPLEMENTARY MATERIALS – PART II

Supplementary materials to: Lenka Bustikova, David Siroky, Saud Alashri, Sultan Alzahrani. Predicting Partisan Responsiveness: A Probabilistic Text Mining Time-Series Approach. *Political Analysis*.

Expert Topic Validation

To address the validity of selecting the topics using experts, we compare the topic selection to two respected datasets of political parties, and then discuss the individual steps in topic selection. Before doing so, we first note that n-grams were ranked based on their frequency and only then did the experts assign topics to top ranked n-grams.

We compare our expert-selected topics with the Manifesto Data and the Chapel Hill Expert Survey, two highly used sources of information about parties. We compared the issues that the experts in our project identified, based on the lists of highly ranked n-grams, with the issues identified by two prominent data sets that were constructed to reduce the complexity of party competition onto different issues. The Manifesto Project is a useful validity check for us because it is based on a textual analysis of party manifestos, and the Chapel Hill Survey is helpful as a validity check because it is based on expert assessments.

As noted in the paper, one important advantage of text mining is its ability to capture smaller, niche parties that may have an outsized impact on party competition, even if their own electoral success is minimal. Both datasets exclude the smaller radical right party (*Slovenská pospolitost'*) and the ethnic party (*Híd*) that are analyzed in our paper, and include only the larger niche parties (*SNS* and *SMK,* one radical right and the other ethnic).

This comparison with the Manifesto Project and the Chapel Hill Expert Survey leads us to believe that our expert assessment captures the overarching issues that dominate debates and structure party politics among parties we study. It is possible of course that other experts might have chosen different labels, but we submit that they would probably capture similar dimensions: issues that revolve around national identity, culture and the role that the Slovak state plays in regulating majority-minority relations, both domestically and in the broader context of international relations.

Here we discuss how precisely we made these comparisons: first for the Party Manifesto data and then for the Chapel Hill Survey data.

Party Manifestos

For 2006, which is in the span of the ten-year time period that we cover (2004-2014), the manifesto project allows researchers to search for topics associated with selected keywords and shows which topics have a higher frequency of keywords. It is important to note that the topics identified by the manifesto project are designed to work across countries yet, at the same time, to capture the axis of contention in particular countries. We searched for two keywords: "identity" and "nation" (in Slovak "ident" and "národ" to capture possible variations of the word) for the two parties available in the project (SMK and SNS). The term "ident" was most frequent keyword identified by the manifesto project that covered the following top three topics (for the Hungarian party): *multiculturalism, federalism* and *political authority.* For the radical right party, the term "ident" was most frequently found in topics related to: *multiculturalism*, *culture and freedom,* and *human rights*.

In the 2006 manifesto of the Hungarian party, the term "národ" was most common in these three topics (*coding schemes* are topics in Manifesto terms): *multiculturalism, freedom* and *human rights,* and *non-economic demographic group* (Hungarians). In the manifesto of the radical right party, the term "národ" was most common in the topics: *national way of life, culture,* and *international (relations)*.

The grid and group issues that the experts identified are very comparable to these general categories that populate party manifestos of two parties. The experts used grid-group theory to highlight group issues that relate to *multiculturalism, national way of life* and *language*.

Grid issues consider ethnicity and nationality in the context of governance, which the manifesto project captures through the coding schemes *international* and *economic issues tied to identity*. The experts for this paper similarly identified three issues: *the EU and EU enlargement, Economics* and *Interstate Relations*. The EU and EU enlargement is an international issue and enlargement is specific to Slovakia, which was once an EU accession country. Interstate relations capture mostly Slovak's contested relationship with the neighboring country. Economic issues are politicized, due to the regional transfers to the Hungarian minority, which is spatially concentrated, and due to the Slovak state's commitments to fund Hungarian minority education from state funds.

The issues that the experts in our project identified as structuring party competition between two political opposites are comparable to the party manifesto project, which codes party platforms and maps issues in manifestos using coding schemes designed to work across cross-national settings.

Chapel Hill Expert Survey

We also validate our choice of grid-group topics against the Chapel Hill Expert Survey of parties. The expert surveys include the radical right SNS and the ethnic Hungarian SMK, just like the Party Manifesto data, and ignores the other two smaller niche parties. In 2014, the most important issues for the radical right party was: *nationalism, ethnic minorities, European integration and anti-elite rhetoric*. For the ethnic party, the top issues were: *ethnic minorities, nationalism and urban vs. rural issues*. The last issue reflects the fact that the Hungarian minority is concentrated in the rural, agricultural part of the country. These topics are mirrored in our selection of issues that relate to national identity and issues of state-minority relations. For the earlier years (2006 and 2010), the expert survey scores the salience of issues for individual parties. For both the radical right party and the Hungarian ethnic party, "*ethnicity*" is the most the salient issue, both in 2006 and 2010. The salience of *the EU* has increased over time for the radical right party. For the ethnic Hungarian party, *the EU* was the second most important issue in the period between 2002-2010.

Discriminative Keywords

keyword	Translation	Weight	Accuracy = 99.7%
'deaktivované'	' deactivated '	1.55055014	
'sns'	' sns '	0.924391191	
'slovenská'	' Slovak '	0.541050848	
'slovenskej'	' Slovakia '	0.491655838	
'reakcie'	' reaction '	0.481781625	
'proti'	' against '	0.478153027	
'počet'	' number '	0.458872542	
'národná'	' national '	0.412188352	
'ľudí'	' people '	0.376510721	
'veľmi'	' very '	0.264727204	
'nás'	' us '	0.23719754	
'no'	'no'	0.199679605	
'im'	' them '	0.185337749	
'ľudia'	' people '	0.181027349	
'slovenských'	' Slovak '	0.173800341	
'organizácie'	' organization '	0.172172491	
'sp'	' sp '	0.141536069	
'kampane'	' Campaign '	0.138867043	
'podujatÍ'	' events '	0.129533743	
'aktuÁlne'	' Currently '	0.129522437	
'kalendÁr'	' Calendar '	0.129522437	
'problém'	' problem '	0.122230131	
'mali'	' should '	0.121582578	
'mu'	' him '	0.114508283	
'meniny'	' Day '	0.113628392	
'2013'	'2013 '	0.113623291	
'strana'	' Party '	0.109512914	
'slovensku'	' Slovakia '	0.106042024	
'zahraničné'	' Foreign '	0.095371724	
'bratislave'	' Bratislava '	0.088566836	
'rafaj'	' Rafaj '	0.080866826	
'človek'	' Human'	0.0793077	
'zaujímavosti'	' things '	0.070696927	
'obete'	' victims '	0.064175609	
'národ'	' nation '	0.061972987	
'naozaj'	' really '	0.060237009	
'dokonca'	' even '	0.059281983	
'národa'	' nation '	0.057780431	
'približne'	' about '	0.05337069	
'usa'	' usa '	0.051350654	
'polícia'	' police '	0.044575067	
'bratislava'	' Bratislava '	0.041444171	

'slovákov'	' Slovaks '	0.040013043
'svete'	' world '	0.038659676
'ľudských'	' Human '	0.034614226
'vás'	' you '	0.032380802
'naše'	' our '	0.030663609
'verejnosť'	' public '	0.029920248
'opäť'	' Again '	0.029432798
'kosovo'	' Kosovo '	0.028388215
'počas'	' during '	0.020395062
'nemzeti'	' nemzeti '	0.019713152
'cigánsky'	' Gypsy '	0.015716031
'tomu'	' this '	0.013561068
'život'	' Life '	0.013455467
'eÚ'	' eu '	0.013121435
'prípade'	' case '	0.009467876
'svojej'	' their '	0.008075208
'szóló'	' szóló '	0.007881191
'slovenského'	' Slovak '	0.007107388
'deti'	' Children '	0.006812746
'cigáni'	' Gypsies '	0.005152523
'veterná'	' wind '	0.004894346
'pripomÍname'	' Recalls '	0.004860942
'výtvarných'	' Fine '	0.004777273
'alkotmánybíróság'	' alkotmánybíróság '	0.004311514
'gallery'	' gallery'	0.003731129
'miliardové'	' billions '	0.003402031
'niekoľko'	' few '	0.002848156
'všetci'	' all '	0.002847161
'pôde'	' soil '	0.002735387
'spôsobila'	' caused '	0.002726734
'musí'	' must '	0.002607129
'vpn'	' vpn '	0.002606161
'čas'	' Time '	0.002561858
'životoch'	' lives '	0.00249813
'vyžiadala'	' requested '	0.002153863
'tzv'	' called '	0.001992805
'opozičná'	' opposition '	0.001987171
'obyvateľstva'	' population '	0.00195962
'Čítať'	' Read '	0.001682242
'végre'	' végre '	0.001611748
'napríklad'	' example '	0.001549526
'niekto'	' someone '	0.001527485
'umelcov'	' artists '	0.00141285
'mal'	' should '	0.001346905
'tatry'	' Mountains '	0.001275455
'smršť'	' whirlwind '	0.001253325

'iniciatíva'	' Initiative '	0.001169738
'ďalej'	' further '	0.001127087
'eddigi'	' eddig '	0.001043424
'polgármester'	' polgármester '	0.000717641
'eugen'	Eugen '	0.000650234
'ember'	' ember '	0.00045086
' már'	' Bier '	0.000386133
'akarja'	' akarja '	0.000362257
'kedy'	' when '	0.000316366
'škody'	' damage '	0.000280989
'eu'	' eu '	0.000278928
'tých'	' the '	0.000262458
'rafael'	'Rafael '	0.000190365
'szükség'	' szükség '	0.000179876
'iných'	' other '	0.000150484
'amennyiben'	' Amennyiben '	-2.95E-06
'poslanci'	' Members '	-4.81E-05
'členovia'	' members '	-0.00024346
'nemzet'	' nemzet '	-0.000406461
'mohli'	' could '	-0.000416907
'szabad'	' szabad '	-0.000440116
'szükséges'	' szükséges '	-0.000456153
'elnök'	' elnök '	-0.000477807
'cigánov'	' Gypsies '	-0.000507851
'felvidéki'	' highland '	-0.000615916
'“'	'	-0.000646417
'slovenska'	' Slovakia '	-0.000671145
'zasiahla'	' hit '	-0.000712208
'anyanyelv'	' anyanyelv '	-0.000713751
'hovorí'	' speaks '	-0.000713777
'Švejna'	' Švejna '	-0.000770229
'szolgáltatók'	' szolgáltatók '	-0.000846135
'mestskej'	' urban '	-0.000860601
'minisztérium'	' Minisztérium '	-0.000967865
'vznikla'	' established '	-0.000999635
'niečo'	' something '	-0.001036259
'gazdaság'	' gazdaság '	-0.001465647
'főleg'	' főleg '	-0.001657283
'tárca'	' Tarco '	-0.001688145
'kérdés'	' kérdés '	-0.001727875
'bastrnák'	' Bastrnák '	-0.001787446
'egyetemi'	' egyetemi '	-0.001874018
'számunkra'	' számunkra '	-0.001887012
'oktatási'	' Oktatási '	-0.001999175
'nikdy'	' Never '	-0.002024058
'jeden'	' one '	-0.002113074

'elsősorban'	' elsősorban '	-0.002145427
'választások'	' választások '	-0.002223819
'voči'	' against '	-0.002388661
'támogatás'	' támogatás '	-0.002545526
'solymos'	' Solymos '	-0.0028394
'egyben'	' egyben '	-0.002943349
'kapcsolattartás'	' kapcsolattartás '	-0.002967461
'utóbbi'	' utóbbi '	-0.003140997
'polgárok'	' polgárok '	-0.0032577
'belső'	' belső '	-0.003268736
'sajnos'	' sajnos '	-0.003764756
'zsolt'	' zsolt '	-0.004103861
'jogi'	' Jogi '	-0.004189328
'magyarság'	' magyarság '	-0.0042248
'fiatal'	' fiatal '	-0.004346372
'munkát'	' Munky '	-0.004712586
'komoly'	' komoly '	-0.004804435
'kisebbségi'	' kisebbségi '	-0.004891099
'slovensko'	' Slovakia '	-0.005074867
'strany'	' Party '	-0.006472302
'gál'	' Gal '	-0.006631466
'pospolitosti'	' Togetherness '	-0.006791191
'eur'	' euro '	-0.006995333
'robert'	'Robert '	-0.007304447
'kettős'	' Kettős '	-0.008221593
'éve'	' éve '	-0.008536573
'szervezet'	' szervezet '	-0.010603855
'megfogalmazott'	' megfogalmazott '	-0.011529994
'község'	' község '	-0.012770578
'gyakran'	' gyakran '	-0.0134391
'psa'	' Dog '	-0.01421877
'parlament'	' Parliament '	-0.014723774
'obce'	' village '	-0.017151084
'jött'	' Jott '	-0.017936391
'megyei'	' megyei '	-0.019226565
'hasonló'	' hasonló '	-0.020063889
'kapott'	' bonnet '	-0.020568249
'idő'	' idő '	-0.021154207
'r7'	' r7 '	-0.021504681
'smeru'	' direction '	-0.021732809
'szorítsunk'	' szorítsunk '	-0.022172932
'oldalra'	' oldalra '	-0.023245153
'képviselői'	' képviselői '	-0.023300631
'hivatal'	' hivatal '	-0.023881016
'miniszter'	' miniszter '	-0.024676648
'farkas'	'Farkas '	-0.025509326

'millió'	' millio '	-0.027257157
'közösség'	' közösség '	-0.027436241
'uniós'	' UNIOS '	-0.028775709
'irányában'	' irányában '	-0.030757012
'jelentős'	' jelentős '	-0.033066063
'János'	'János '	-0.03320563
'állam'	' állam '	-0.034340965
'kis'	' kis '	-0.034871872
'bizonyára'	' bizonyára '	-0.035824096
'unió'	' unió '	-0.042581715
'például'	' például '	-0.0450093
'percenta'	' percent '	-0.045666775
'szeretnék'	' szeretnék '	-0.045821608
'jövő'	' jövő '	-0.048409513
'túl'	' quiver '	-0.050208683
'vláda'	' Government '	-0.053350742
'szociális'	' szociális '	-0.054139319
'gábor'	' Gabor '	-0.057858612
'béla'	' Bela '	-0.058068268
'melyik'	' melyik '	-0.062545246
'bugár'	' Bugár '	-0.065458546
'zákona'	' Law '	-0.065848248
'költségvetés'	' Költségvetés '	-0.066973647
'kulturális'	' CULTURAL '	-0.067663713
'nr'	' NR'	-0.068960844
'lengyel'	' Lengyel '	-0.07075565
'jobb'	' jobb '	-0.074513576
'iván'	'Iván '	-0.076057279
'fico'	' fico '	-0.078015373
' okt'	' Oct '	-0.081066282
'elnöke'	' elnök '	-0.085204177
'alajos'	'Alajos '	-0.085615097
'poslanec'	' MP '	-0.093226732
'vlády'	' Government '	-0.095600191
'Árpád'	'Árpád '	-0.098433269
'teljesen'	' teljesen '	-0.098484333
'választási'	' Választási '	-0.100696499
'pártja'	' Pártja '	-0.103463637
'szlovákiában'	' szlovákiában '	-0.105658417
'psov'	' dogs '	-0.109376039
'sikerült'	' sikerült '	-0.111266223
'mészáros'	' Meszaros '	-0.11246429
'csupán'	' csupán '	-0.115277423
'általa'	' Alta '	-0.120015361
'ezer'	' ezer '	-0.122474032
'pál'	' Pal '	-0.135604978

'simon'	'Simon '	-0.151973719
'orbán'	'Orbán '	-0.152483306
'pénzügyi'	' pénzügyi '	-0.167138628
'alelnöke'	' alelnöke '	-0.191895483
'csáky'	' Csáky '	-0.204880357
'lászló'	' Laszlo '	-0.20543531
'mai'	' mai '	-0.209767942
'józsef'	'József '	-0.214136358
'éves'	' eves '	-0.230591982
'törvény'	' törvény '	-0.231745297
'emberek'	' emberek '	-0.246543913
'tags'	' tags '	-0.254677632
'helyi'	' helyi '	-0.262754371
'smer'	' direction '	-0.268468903
'Érsek'	' Érsek '	-0.286703907
'szlovákiai'	' szlovákiai '	-0.298332149
'szlovákia'	' Szlovákia '	-0.348550495
'parlamenti'	' Parlamenti '	-0.359439897
'gazdasági'	' gazdasági '	-0.379334411
'szlovák'	' szlovák '	-0.65036627
'magyar'	' magyar '	-0.849083361
'mkp'	' MLP '	-0.860817463
'híd'	' Hid '	-1.145614012

Expert selection of topics

Topic 1 - *EU*: Europe, Enlargement

Topic 2 - *Minorities*: (Hungarians and Gypsies/Roma): Gypsies, Roma, Hungarians, Gypsy problem, minority rights, rights of Hungarians, dual citizenship

Topic 3 - *Identity*: people, Slovaks, nation (narod, ludi, ludia, Slovaci, Slovensky)

Topic 4 - *Language*: language law, law, education, school, minority language (zakon, skolstva, skolstvo, skoly), linguistic, school act

Topic 5 - *Economics*: Regional development, grant, regions, economic policy

Topic 6 - *Interstate Relations*: Hungary (Madarsko), Orban, Budapest

Top Keywords and N-Grams

Terms	Frequency	Translation	Frequency
	4239		4239
proti		against	
	3960		3960
strana		page	
	3638		3638
sns		sns	
	3435		3435
slovenských		Slovak	
	3290		3290
organizácie		organization	
	3285		3285
bratislave		bratislava	
	3064		3064
verejnosť		public	
	3042		3042
ľudských		human	
	3035		3035
obete		victims	
	2993		2993
slovenskej		Slovakia	
	2963		2963
vznikla		created	
	2915		2915
zväzu		Union	
mestskej		urban	
	2914		2914
iniciatíva		initiative	
	2909		2909
pôde		soil	
	2890		2890
zasiahla		hit	
	2864		2864
spôsobila		caused	
	2862		2862
umelcov		artists	
vpn		vpn	
	2855		2855
tatry		Tatras	
	2853		2853
opozičná		opposition	
	2852		2852

životoch		lives	
vyžiadala		requested	
	2847		2847
miliardové		billions	
	2844		2844
výtvarných		Fine	
verejnosť proti		public against	
	2843		2843
ľudských životoch		human life	
mestskej organizácie		urban organizations	
smršť		whirlwind	
	2842		2842
slovenských výtvarných		Slovak Fine	
zasiahla veterná smršť		hit windstorm	
výtvarných umelcov		artists	
umelcov vznikla		established artists	
vznikla opozičná		established opposition	
miliardové škody.kalendár		billions škody.kalendár	
pôde mestskej organizácie		organization of urban land	
veterná smršť vyžiadala		windstorm requested	
vznikla opozičná iniciatíva		opposition arose Initiative	
kampanepripomíname si19		kampanepripomíname si19	
spôsobila miliardové		caused billions	
tatry zasiahla veterná		Tatra hit the wind	
škody.kalendár		škody.kalendár	
mestskej organizácie zväzu		Urban union organizations	
iniciatíva verejnosť proti		Initiative against public	
kampanepripomíname		kampanepripomíname	
zväzu slovenských výtvarných		Slovak Union of Visual	
opozičná iniciatíva		opposition initiative	
iniciatíva verejnosť		public initiative	
zasiahla veterná		hit the wind	
spôsobila miliardové škody.kalendár		caused billions škody.kalendár	
umelcov vznikla opozičná		artists created opposition	
veterná smršť		windstorm	
pôde mestskej		urban land	
organizácie zväzu slovenských		Association of Slovak organizations	
zväzu slovenských		Slovak Association	
si19		si19	
organizácie zväzu		union organizations	
výtvarných umelcov vznikla		artists created	
tatry zasiahla		Mountains hit	
veterná		wind	
smršť vyžiadala		storm requested	
opozičná iniciatíva verejnosť		public opposition initiative	
slovenských výtvarných umelcov		Slovak artists	

strany	2767	Party	2767
slovensku	2490	Slovakia	2490
predseda	2111	chairman	2111
meniny	1922	name-day	1922
ľudí	1905	people	1905
národnej	1887	national	1887
nás	1816	us	1816
slovenska	1767	Slovakia	1767
mali	1751	mali	1751
slovensko	1749	Slovakia	1749
vlády	1741	Government	1741
veľmi	1716	very	1716
tomu	1634	it	1634
ide	1539	regards	1539
tohto	1519	this	1519
mal	1483	had	1483
ďalej	1469	further	1469
povedal	1450	said	1450
im	1449	them	1449
vláda	1431	government	1431
ľudia	1386	people	1386
republiky	1357	Republic	1357
slovenského	1314	Slovak	1314

	1290		1290
prípade		case	
	1287		1287
štátu		State	
	1264		1264
ktorej		which	
	1234		1234
občanov		citizens	
predsedu		President	
	1224		1224
zákona		Act	
	1204		1204
problém		problem	
	1175		1175
rady		Board	
	1161		1161
jeden		one	
no		yet	
	1156		1156
napríklad		for example	
	1141		1141
parlamentu		Parliament	
	1130		1130
nr		NR	
	1126		1126
neho		it	
	1115		1115
chce		wants	
	1113		1113
mu		him	
podpredseda		vice-chairman	
	1103		1103
národa		nation	
	1101		1101
	2010		2010
	1093		1093
voči		towards	
ktorého		whose	
	1092		1092
treba		be	
	1090		1090
slovenskej národnej		Slovak national	
	1070		1070
štát		state	
	1061		1061
slovenskej republiky		Slovak Republic	

slovenský	1057	Slovak	1057
mala	1045	had	1045
tých	1036	those	1036
všetci	1030	all	1030
jána	1029	John	1029
zákon	1024	law	1024
dokonca		even	
počas	1020	during	1020
naše	1017	ours	1017
predseda sns	1015	President sns	1015
napriek	1010	despite	1010
totiž	1000	namely	1000