Data Analysis and Initial Results

Fabian Aiolfi

2023-12-02

Introduction

This document provides an overview of the current status of the data analysis and initial results. It starts with an overview of all eleven CEE EU member states. Descriptive statistics display insights into the available data. I then continue by modelling emigration on incumbent vote change using a fixed effects model. Results show that an increase in emigration is associated with a decrease in incumbent vote share.

In a second step I focus on the effects of service cuts in Poland. I again begin with descriptive statistics that provide an overview. Then I model the effects of service cuts on incumbent vote change. Results partially align with what my theory predicts. An increase in the number of children per school and an increase in the number of people per third place is associated with a decrease in incumbent vote share.

CEE EU Member States

Anti-Incumbent Voting: Was the Party Incumbent?

11 CEE EU Member States, NUTS2/3, 1994–2019 (N = 3240)
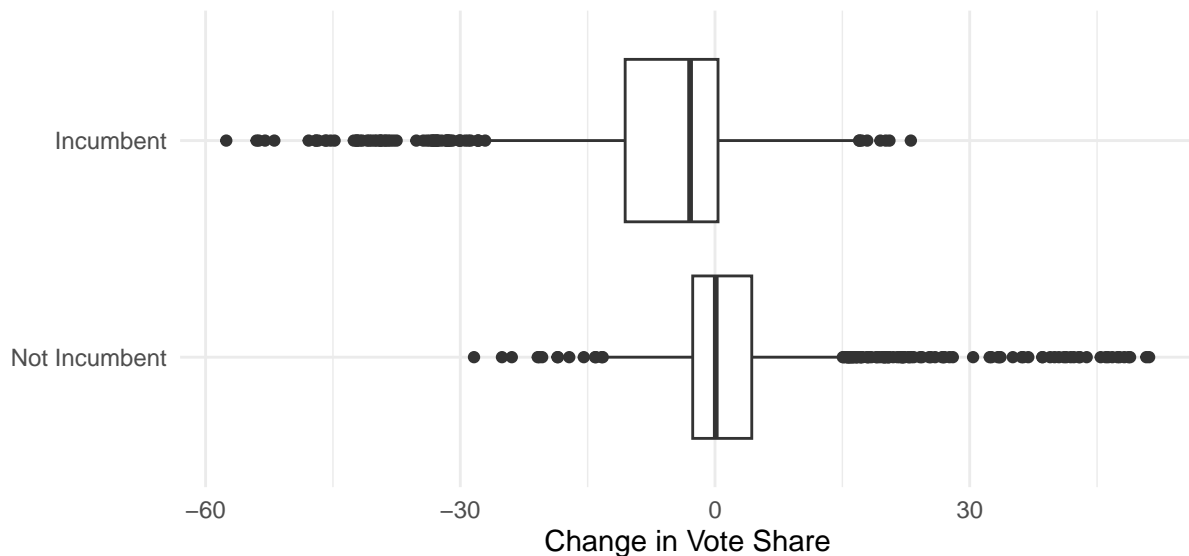
Based on this box plot, we can observe that non-incumbent parties tend to increase their vote share compared to the previous election. On the other hand, incumbent parties seem to get punished by the electorate and more often see a decrease in their vote share. This assumption is supported by a t-test, performed below. The t-test displays a statistically significant result for the difference in means.

```
##
##  Welch Two Sample t-test
##
## data:  group_true and group_false
## t = -19.586, df = 2539.2, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -8.082278 -6.611203
## sample estimates:
## mean of x mean of y
## -5.432212  1.914528
```

External Emigration

Crude Emigration per 1000 Population at NUTS3 Level

This is an overview of the available external emigration data at NUTS3 level (Poland: NUTS2; Slovenia: NUTS1). Please note that the y-axis varies by country, the x-axis however is the same for all countries. This makes a comparison of available data across different countries easier. Of the eleven CEE EU member states, only ten are displayed as Czechia does not provide data solely on external emigration. Each grey line represents a NUTS region. The red line displays the overall trend over all available years and regions within a country.
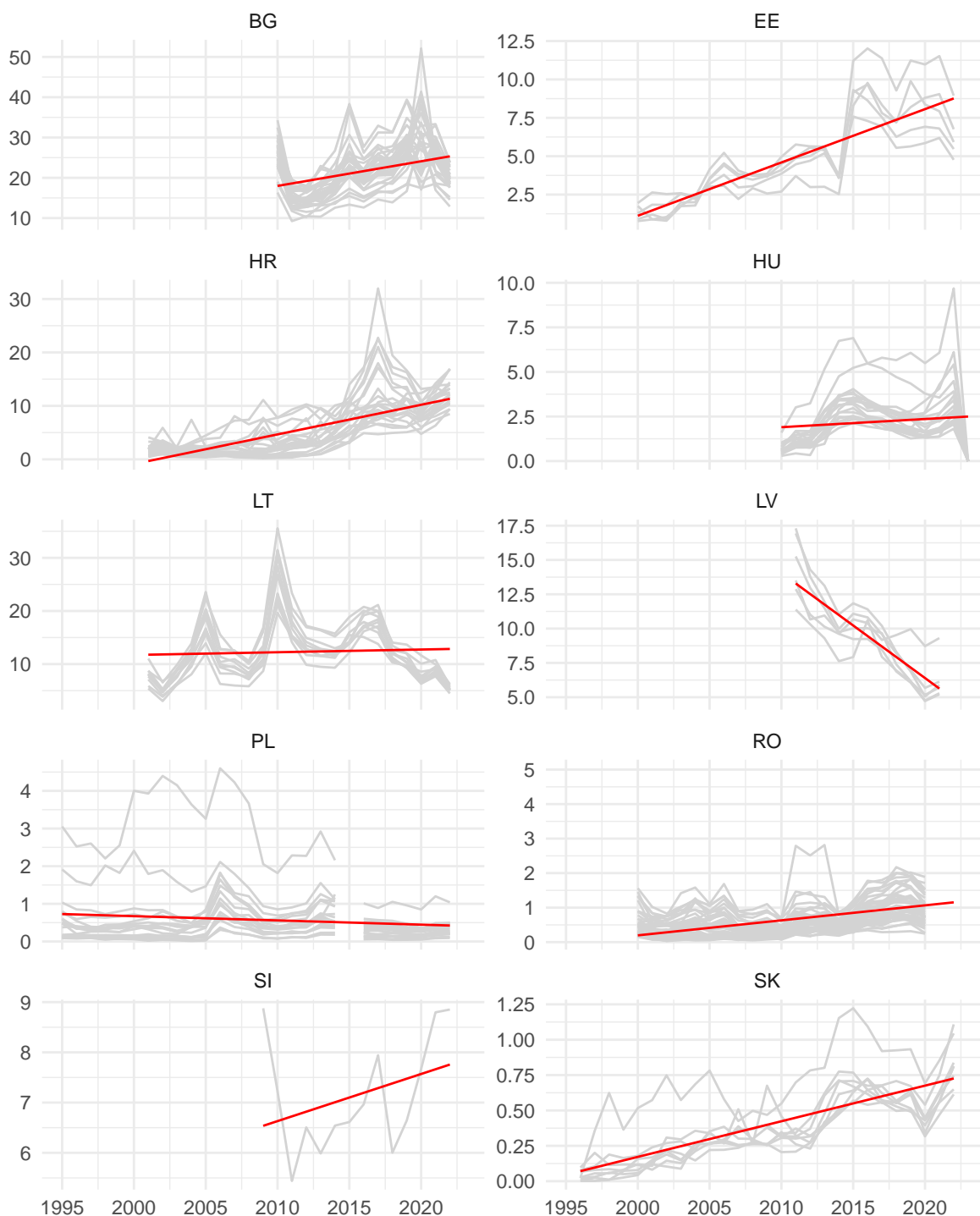
	(1)
average_emigration	$-0.311+$
	(0.170)
Num.Obs.	740
R2 Adj.	0.534
R2 Within Adj.	0.001

$+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001$

Model: Emigration and Incumbent Vote Change

I model emigration rates on incumbent vote change across all NUTS3 regions in CEE EU member states. First, I calculate a region's rolling emigration rate average, taking the average from two consecutive years. E.g., a region's average emigration rate in the year 2010 is the average of the year 2008 and 2009. I do this to smooth out potential outliers and attempt to model a person's perception of emigration in a region.

I use the NUTS3 region and the year as fixed effects when building the model, accounting for region- and time-specific factors that may confound the results.

The coefficient is negative, which shows that an increase in the average emigration rate in a NUTS3 region is associated with a decrease in the change of vote share for the incumbent party. This aligns with my theory. The result is statistically significant at the $p < 0.1$ level. The standard errors are clustered at the NUTS3 level, as observations within a region are more similar to each other than to observations in other regions. The overall model fit is 53.34%. However, the "within" model fit is very low (0.24%), which indicates that average emigration does not explain variation in the dependent variable within each fixed-effects group.

	(1)
average_emigration	−0.250
	(0.167)
lrgen_fctCentre Left	−3.251**
	(1.193)
lrgen_fctCentre Right	−3.155**
	(1.179)
lrgen_fctLeft	−14.360***
	(1.850)
lrgen_fctRight	−3.411*
	(1.324)
Num.Obs.	740
R2 Adj.	0.553
R2 Within Adj.	0.042

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Model: Emigration, Incumbent Vote Change and Party Orientation

As a small excursion, I use data from the Chapel Hill Expert survey to categorise parties by their overall ideological stance: Left, Centre Left, Centre, Centre Right and Right. I add this new information to the model, which uses Centre as the reference category.

Average emigration still shows a negative coefficient and is borderline statistically significant. All party coefficients are statistically significant and the overall model fit as well as the "within" model fit have increased. We can also observe that Left parties fare far worse than other parties when emigration increases.

Service Cuts in Poland

Population Statistics

I now focus on service cuts in Poland, starting with the country's overall population and primary school population. NUTS2 data is used, because Poland only provides data at this level. Each line represents a Polish NUTS2 region, of which there are 16. A plot's overall trend is displayed with a red trendline.

Contrary to my expectations, the overall population has not decreased in the past 20 years. However, the number of primary aged school children, aged seven to twelve, has decreased.

Entire Population in Poland by NUTS2
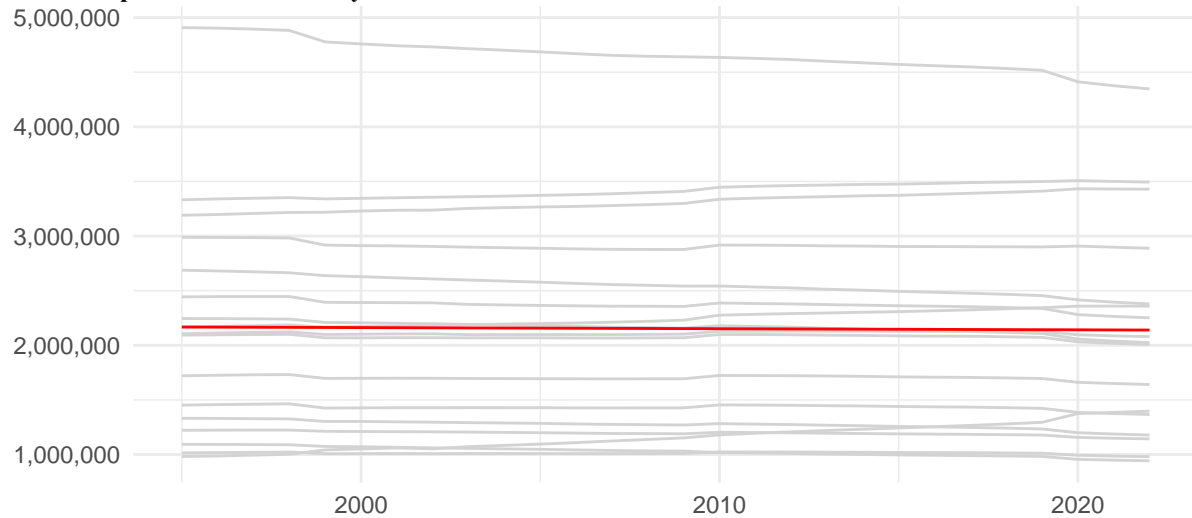
**Primary Age Population in Poland by NUTS2**
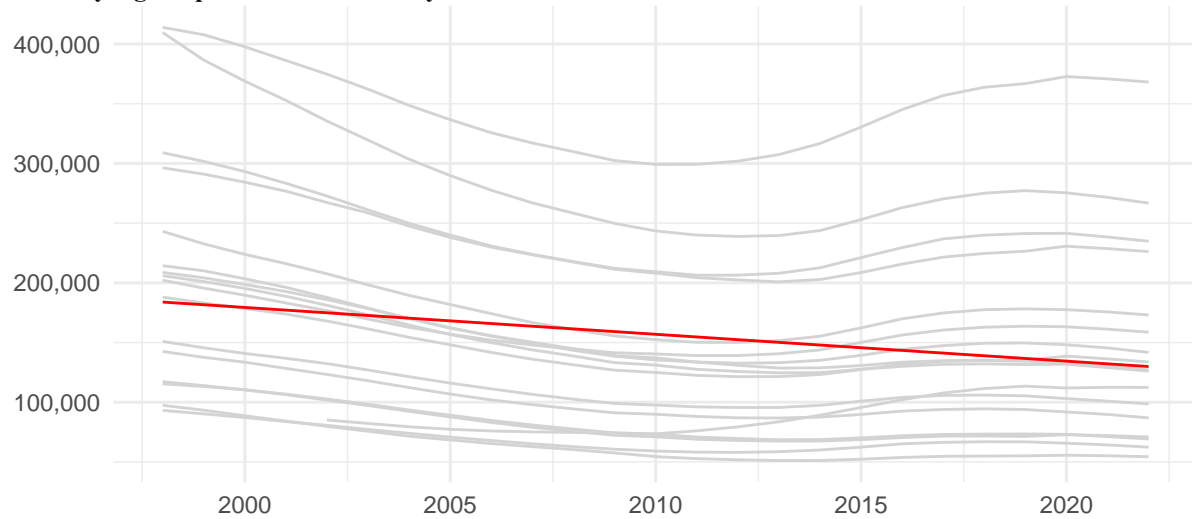
Ratios

Also contrary to my expectations, all services that I investigate generally see a decrease in their ratios. This means that over time, there are less people per institution in a NUTS2 region. E.g., less children in a school or less citizens per hospital.
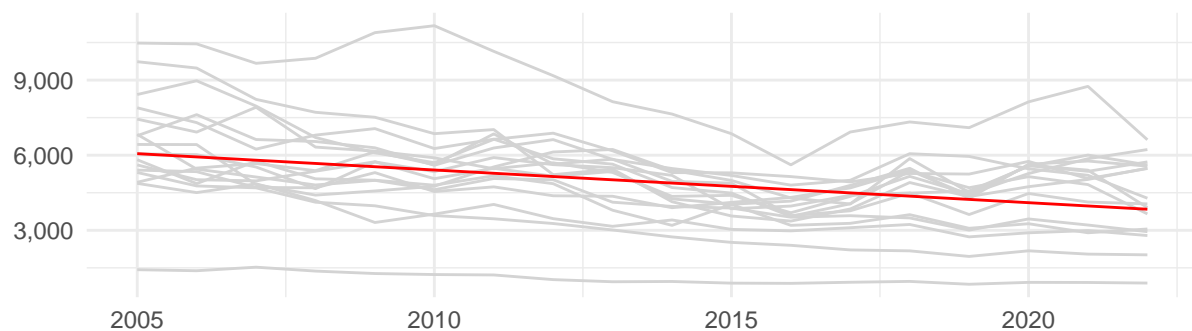
Number of Primary School Children per School by NUTS2 in Poland
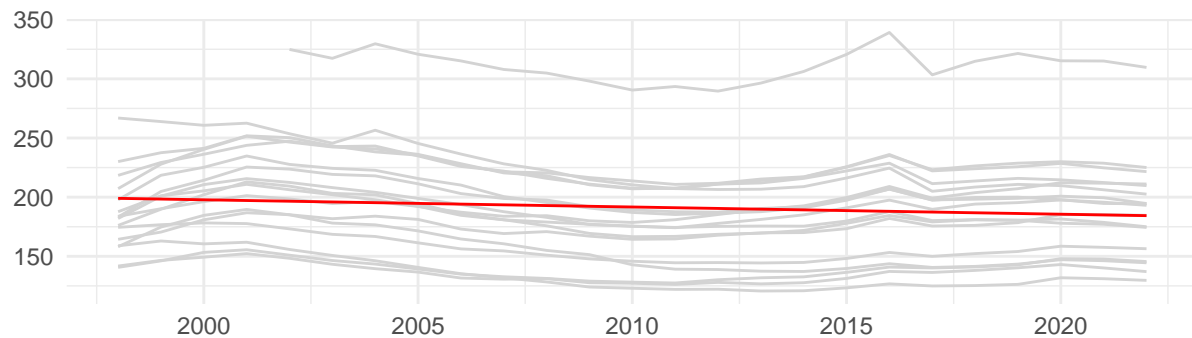
Number of People per Hospital by NUTS2 in Poland
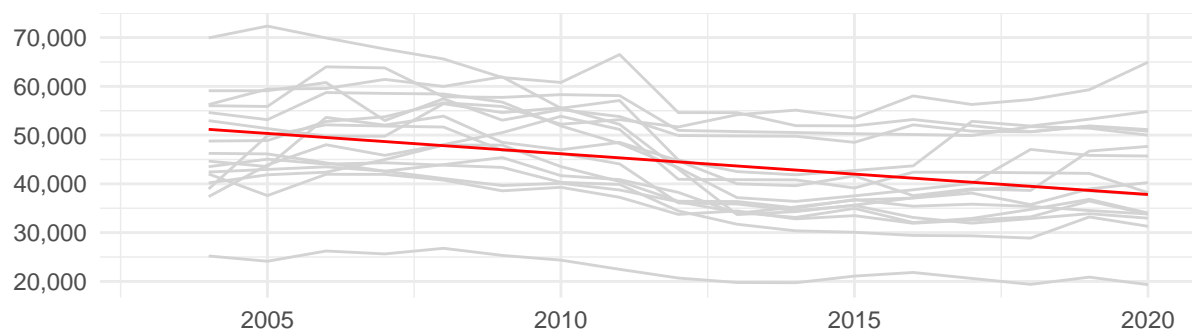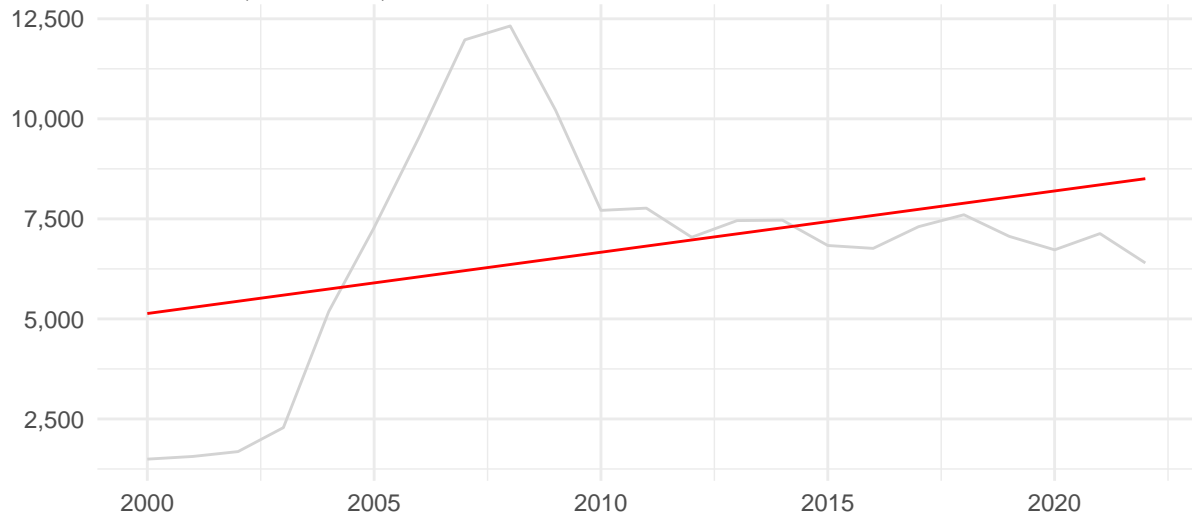

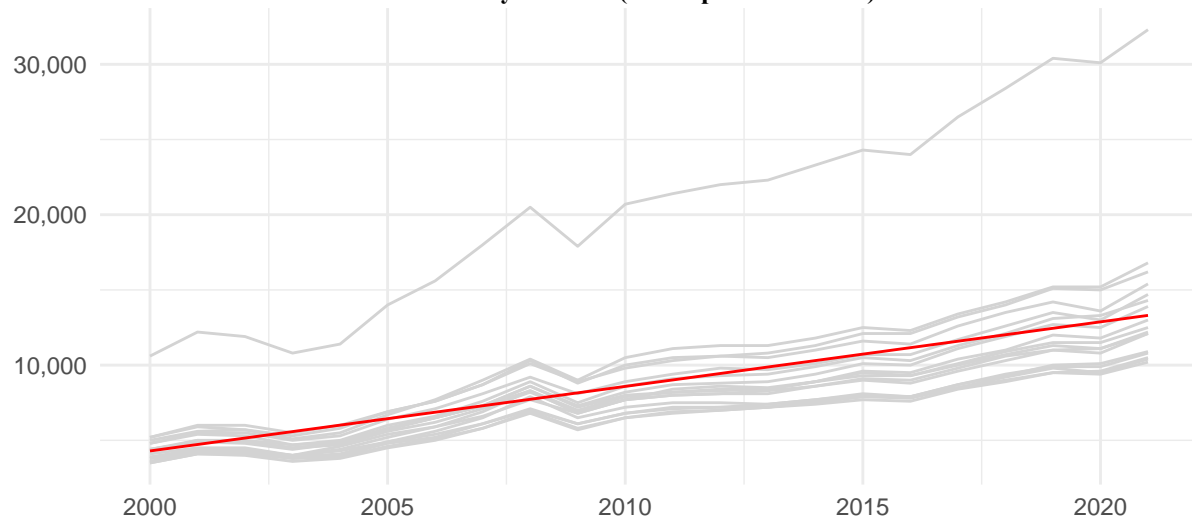**Number of People per Third Place by NUTS2 in Poland**

Control Variables

Both remittances and GDP show an increase over time. I use these data as control variables in my models. Data on remittances are only available at the national level.

Remittance Inflows (US$ million) for all of Poland

GDP at Current Market Prices in Poland by NUTS 2 (Euros per Inhabitant)

	(1)	(2)	(3)
ratio_schools	−0.089		
	(0.066)		
emigration_election_year_per_1000	0.128	−0.055	−0.152
	(0.110)	(0.168)	(0.154)
gdp	0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)
ratio_hospitals_all_population		0.000	
		(0.000)	
ratio_third_places			−0.001*
			(0.000)
Num.Obs.	186	186	186
R2	0.502	0.502	0.503
R2 Adj.	0.438	0.438	0.439
R2 Within	0.001	0.001	0.003
R2 Within Adj.	−0.017	−0.017	−0.016
AIC	1439.4	1439.5	1439.1
BIC	1510.4	1510.4	1510.1
RMSE	10.30	10.30	10.29
Std.Errors	by: nuts2016	by: nuts2016	by: nuts2016
FE: nuts2016	X	X	X
FE: year	X	X	X

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

LM Models: Anti-Incumbent Voting in Poland at NUTS2 Level

I first model separate ratios on incumbent vote change while controlling for emigration and GDP. Models 1, 2 and 3 each use a ratio separately. NUTS2 regions and years are used as fixed effects in all models. The emigration rate is the rolling emigration average between two national elections. Remittances are removed due to collinearity.

Some coefficients point in the right direction but overall no results are statistically significant at conventional levels. An increase in the number of children per school is associated with a decrease in the incumbent vote share. However the p-value is high ($p = 0.197$). Both hospital and third places ratios have coefficients of 0, displaying no effect on incumbent vote share.

	(1)	(2)	(3)	(4)
ratio_schools	−0.109+	−0.092		−0.112*
	(0.055)	(0.054)		(0.048)
ratio_hospitals_all_population	0.000		0.000	0.000
	(0.000)		(0.000)	(0.000)
emigration_election_year_per_1000	0.063	−0.038	−0.218	−0.100
	(0.123)	(0.129)	(0.164)	(0.135)
gdp	0.000	0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)
ratio_third_places		−0.001+	−0.001*	−0.001+
		(0.000)	(0.000)	(0.000)
Num.Obs.	186	186	186	186
R2	0.503	0.503	0.503	0.504
R2 Adj.	0.436	0.436	0.436	0.434
R2 Within	0.002	0.004	0.003	0.005
R2 Within Adj.	−0.022	−0.021	−0.021	−0.026
AIC	1441.2	1440.9	1441.0	1442.7
BIC	1515.4	1515.1	1515.2	1520.1
RMSE	10.29	10.29	10.29	10.28
Std.Errors	by: nuts2016	by: nuts2016	by: nuts2016	by: nuts2016
FE: nuts2016	X	X	X	X
FE: year	X	X	X	X

$+ p < 0.1$, $* p < 0.05$, $** p < 0.01$, $*** p < 0.001$

In a second step, I use different combinations of ratios (models 1, 2 and 3) and finally combine all ratios in model 4. Again, NUTS2 regions and years are used as fixed effects. Both school ratio and third places ratio display statistically significant and negative coefficients, which align with my theory. This means that an increase in the number of children per school and an increase of the number of people per third place is associated with a decrease in the incumbent vote share, while controlling for emigration and GDP. The third ratio, number of people per hospital, still shows no effect.

Sources

…