Data Analysis

Fabian Aiolfi

2023-12-24

Contents

1 Summary **2**

2 Overview of all CEE EU Member States **3**
- 2.1 Anti-Incumbent Voting in all CEE EU Member States . . . 3
- 2.2 External Emigration . . . 4

3 Service Cuts in Poland **8**
- 3.1 Population and Emigration . . . 8
- 3.2 Number of Institutions . . . 11
- 3.3 Ratios of People per Institution . . . 15
- 3.4 Control Variables . . . 19
- 3.5 Modelling Anti-Incumbent Voting in Poland at NUTS2 Level using Ratios . . . 20
- 3.6 Modelling Anti-Incumbent Voting in Poland at NUTS2 Level using Number of Institutions . . . 22
- 3.7 Modelling with Truncated Data . . . 23

4 Service Cuts in Romania **25**
- 4.1 Population and Emigration . . . 25
- 4.2 Number of Institutions . . . 28
- 4.3 Ratios of People per Institution . . . 31
- 4.4 Control Variables . . . 35
- 4.5 Modelling Anti-Incumbent Voting in Romania at NUTS3 Level using Ratios . . . 36
- 4.6 Modelling Anti-Incumbent Voting in Romania at NUTS3 Level using Number of Institutions . . . 38
- 4.7 Modelling with Truncated Data . . . 39

5 Sources **41**

1 Summary

This document provides an overview of my current data analysis. It starts with an overview of all eleven CEE EU member states. Descriptive statistics display insights into the available data on anti-incumbent voting and emigration rates. 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.

I then focus on the effects of service cuts in Poland and Romania. I examine these two CEE EU member states separately. I again begin with descriptive statistics that provide an overview of the available data: population numbers, emigration rates, number of institutions per region and calculated ratios, e.g., number of children per school. Then I model the effects of service cuts on incumbent vote change, again using fixed effects models. I combine different independent variables in order to measure their combined effects. I also use the actual number of institutions instead of the ratio as independent variables, to see if results change. Finally, I truncate the data, removing outliers in an attempt to build models with higher validity.

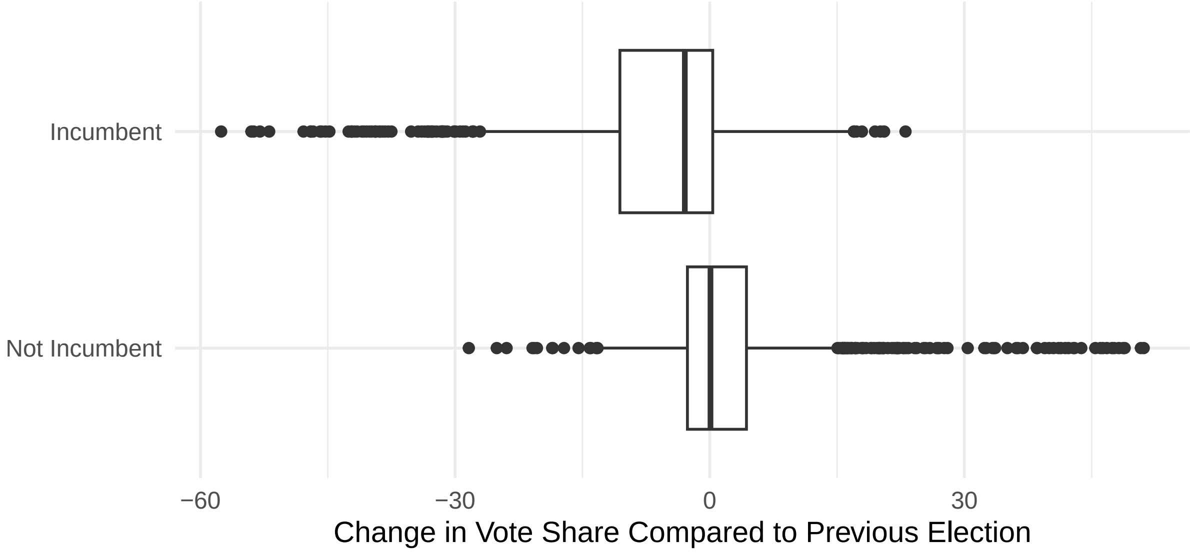
Results partially align with what my theory predicts. In Poland, an increase in the number of children per school and an increase in the number of people per third place is associated with a decrease in incumbent vote share at NUTS2 level. In Romania, a decrease in the number of schools is associated with a decrease in incumbent vote share at NUTS3 level.

Next steps could be to take a region's characteristic into account. E.g., differentiate between rural and urban regions, or regions that border neighbouring countries, in an attempt to group regions with similar trends. Another next step could be to take an incumbent party's orientation into account, as this may have an effect on service cuts. I may also have to find an alternative for the remittances variable, which I include in my models to control for electoral volatility. Remittances are only available for an entire country and are automatically removed due to collinearity when creating fixed effects models.

2 Overview of all CEE EU Member States

2.1 Anti-Incumbent Voting in all CEE EU Member States

Is there a difference in vote share change between incumbent and non-incumbent parties? Overview of 11 CEE EU Member States, NUTS2/3 regions, 1994–2019 ($N = 3240$).



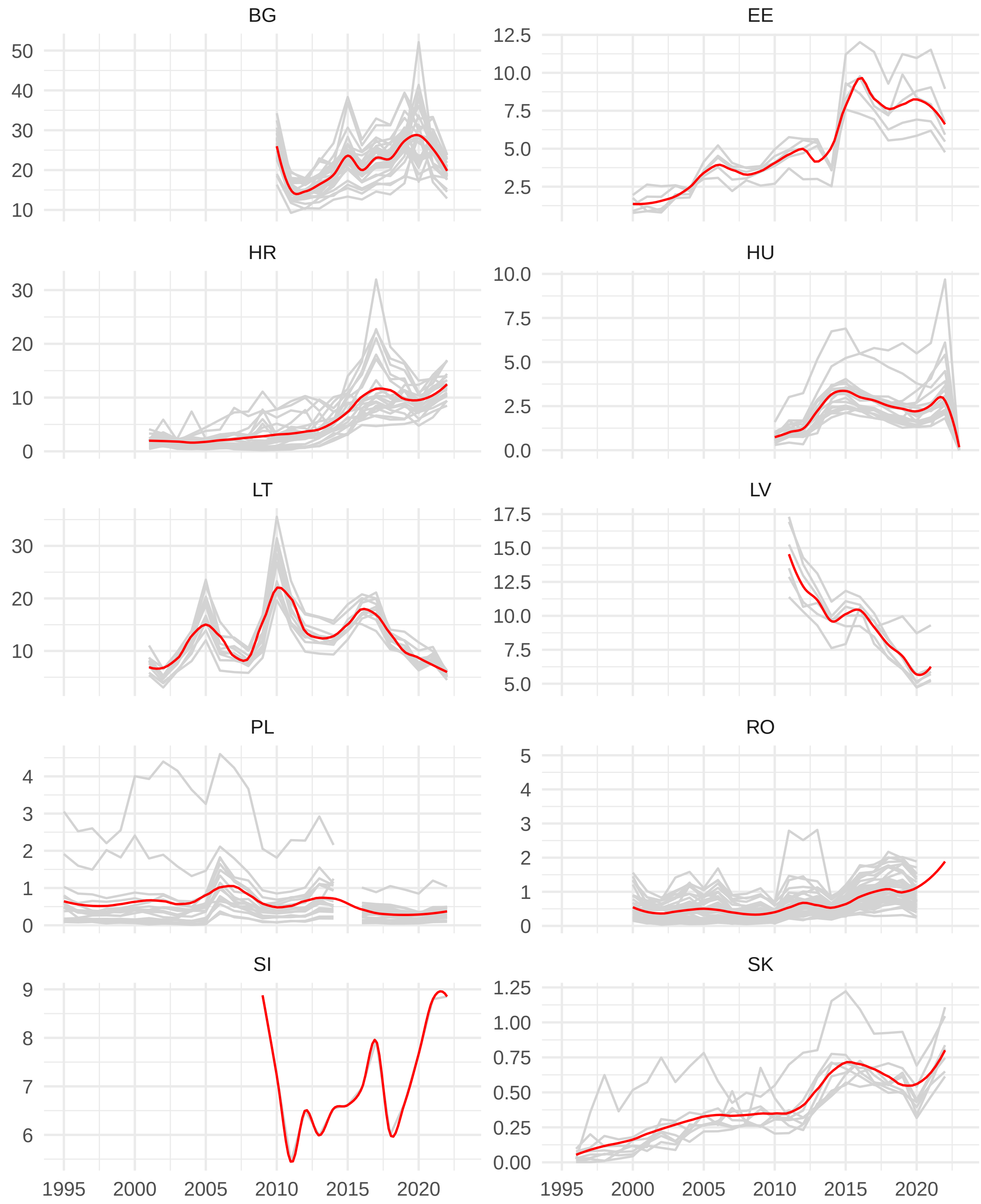
Based on this box plot, we can observe that non-incumbent parties tend to increase their vote share compared to the previous election. On the other hand, incumbent parties seem to get punished by the electorate and more often see a decrease in their vote share. This assumption is supported by a t-test, 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
```

2.2 External Emigration

2.2.1 Crude Emigration per 1000 Population at NUTS3 Level in all CEE EU Member States

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, which applies smoothing to enhance readability, displays the overall trend over all available years and regions within a country.



2.2.2 Emigration and Incumbent Vote Change in all CEE EU Member States

I examine if there is a relationship between emigration rates and an incumbent's change in vote share 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.

This plot provides a basic overview of the relationship between the average emigration rate and the incumbent's change in vote share. There generally seems to be a positive correlation between the two: An increase in average emigration is associated with higher levels of vote change for incumbent parties. This indicates that incumbent parties achieve better results when emigration increases.

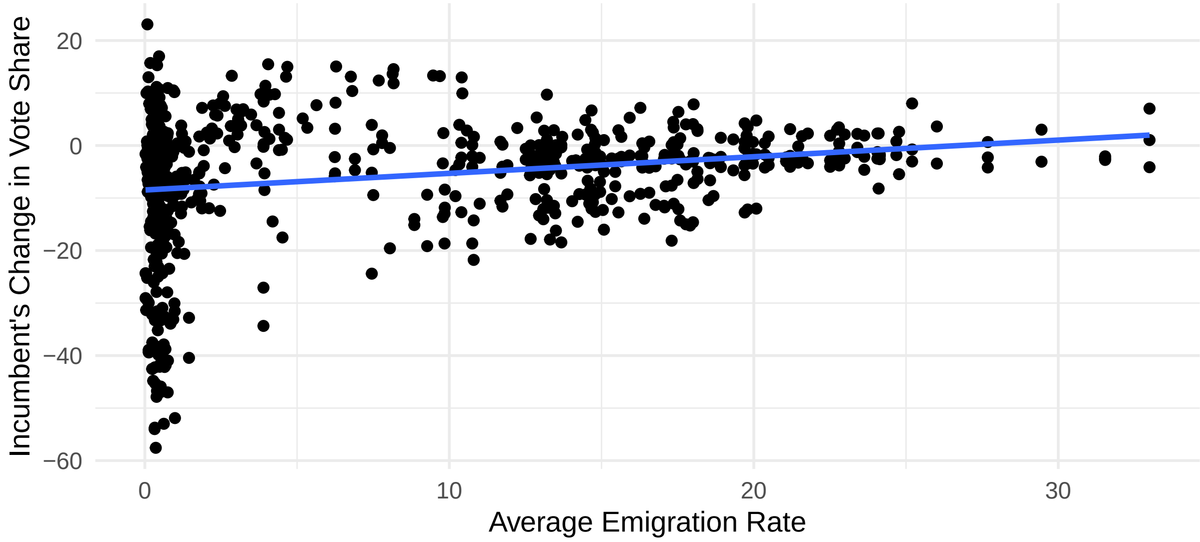


Table 1: DV: Incumbent's Change in Vote Share (all CEE EU MS)

	(1)
Average Emigration	−0.311+
Num.Obs.	740
R2 Adj.	0.534
R2 Within Adj.	0.001
FE: nuts2016	X
FE: year	X

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

In a next step, I build a model using NUTS3 regions and years as fixed effects, accounting for region- and time-specific factors that may confound the results. Results are displayed in Table 1. The average emigration 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. 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 (0.1%), which indicates that average emigration does not explain variation in the dependent variable within each fixed-effects group.

Table 2: DV: Incumbent's Change in Vote Share (all CEE EU MS)

	(1)
Average Emigration	−0.250
Left Party	−14.360***
Centre Left Party	−3.251**
Centre Right Party	−3.155**
Right Party	−3.411*
Num.Obs.	740
R2 Adj.	0.553
R2 Within Adj.	0.042
FE: nuts2016	X
FE: year	X

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Reference Category: Centre Party

2.2.3 Emigration, Incumbent Vote Change and Party Orientation in all CEE EU Member States

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.

As Table 2 displays, average emigration still shows a negative coefficient and is borderline statistically significant ($p = 0.14$). All party coefficients are statistically significant and the overall model fit as well as the "within" model fit have increased compared to the previous model. We can observe that, compared to Centre parties, all other parties suffer a decrease in vote share when emigration increases. Incumbent Left parties see the largest decrease in vote share.

3 Service Cuts in Poland

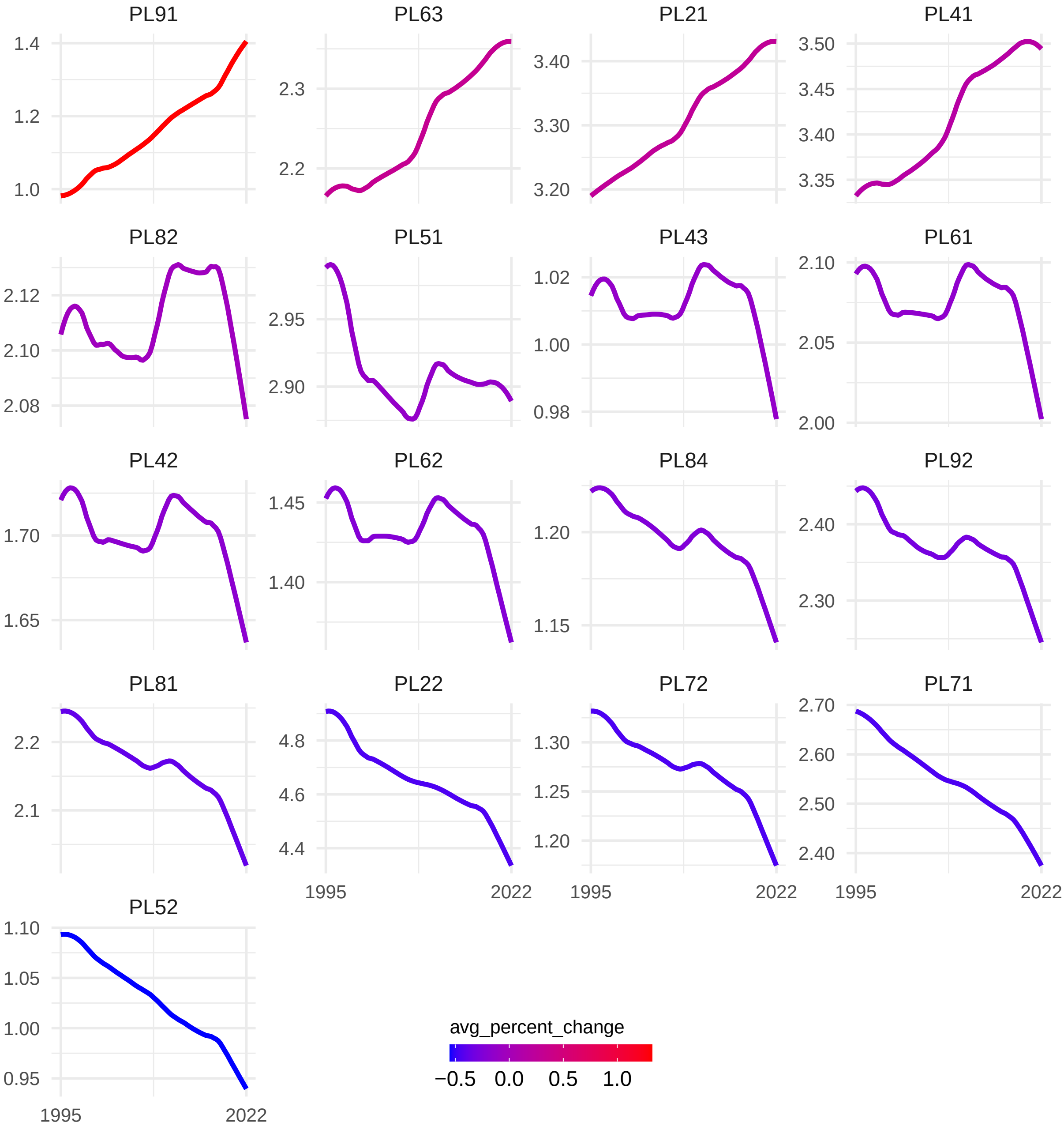
3.1 Population and Emigration

NUTS2 data is used, because Poland only provides data at this level.

3.1.1 Population in Millions in Poland at NUTS2 Level

The population has decreased in a majority of Polish NUTS2 regions. The four regions that have seen the largest increase are regions that include major metropolitan areas, such as the Warsaw Metropolitan Area (PL91), the Pomeranian Voivodeship (PL63, which includes the city of Gdańsk), Lesser Poland Voivodeship (PL21, including Kraków) and Greater Poland Voivodeship (PL41, including Poznań).

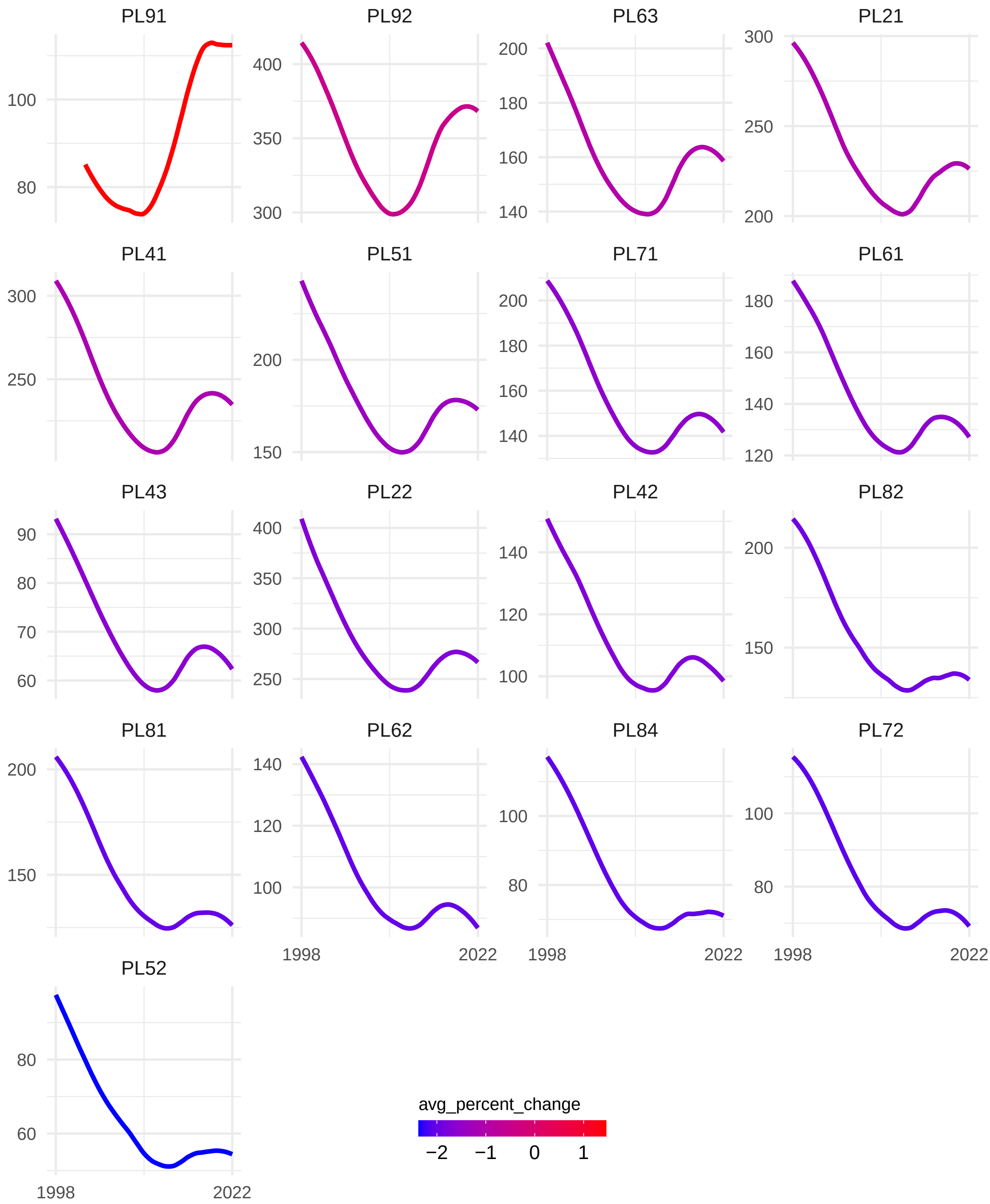
Curves apply smoothing to enhance readability. Plots are ordered by average annual growth rate, with the highest at the top left and decreasing to the right. The growth rate is also indicated by the curve's colour.



3.1.2 Primary Age Population in 1000s in Poland at NUTS2 Level

The primary age population, i.e., children aged seven to twelve, has seen a very similar development across regions apart from Warsaw Metropolitan Area (PL91). This segment of the population decreased steadily until around the year 2010–2012, after which an uptick is visible.

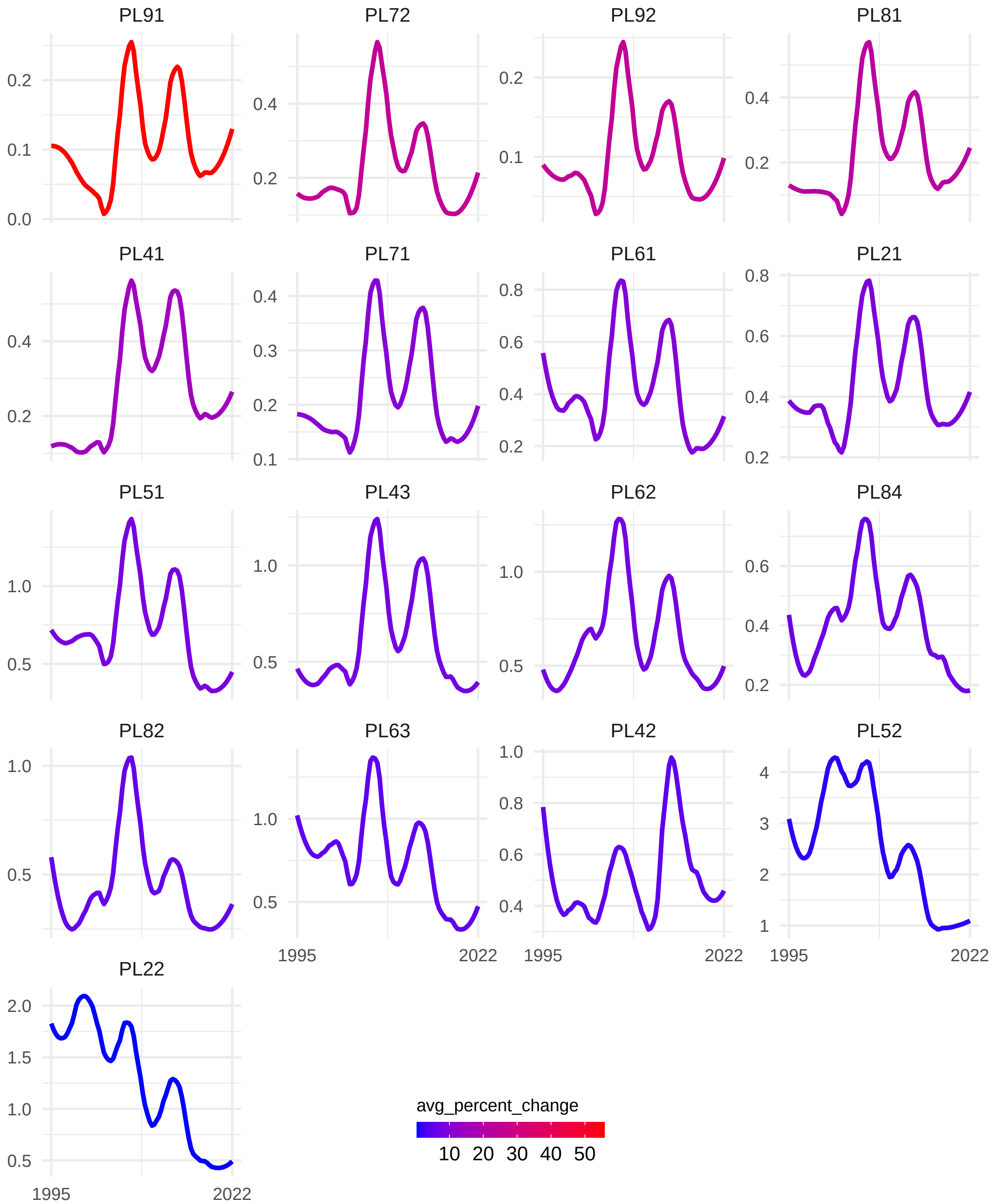
Curves apply smoothing to enhance readability. Plots are ordered by average annual growth rate, with the highest at the top left and decreasing to the right. The growth rate is also indicated by the curve's colour.



3.1.3 External Emigration per 1000 Population from Poland at NUTS2 Level

Most regions display two peaks in external emigration around the years 2006 and 2013/2014. The first peak may be due to Poland's accession to the EU in the year 2004. The reason for the second peak is unclear.

Curves apply smoothing to enhance readability. Plots are ordered by average annual growth rate, with the highest at the top left and decreasing to the right. The growth rate is also indicated by the curve's colour.



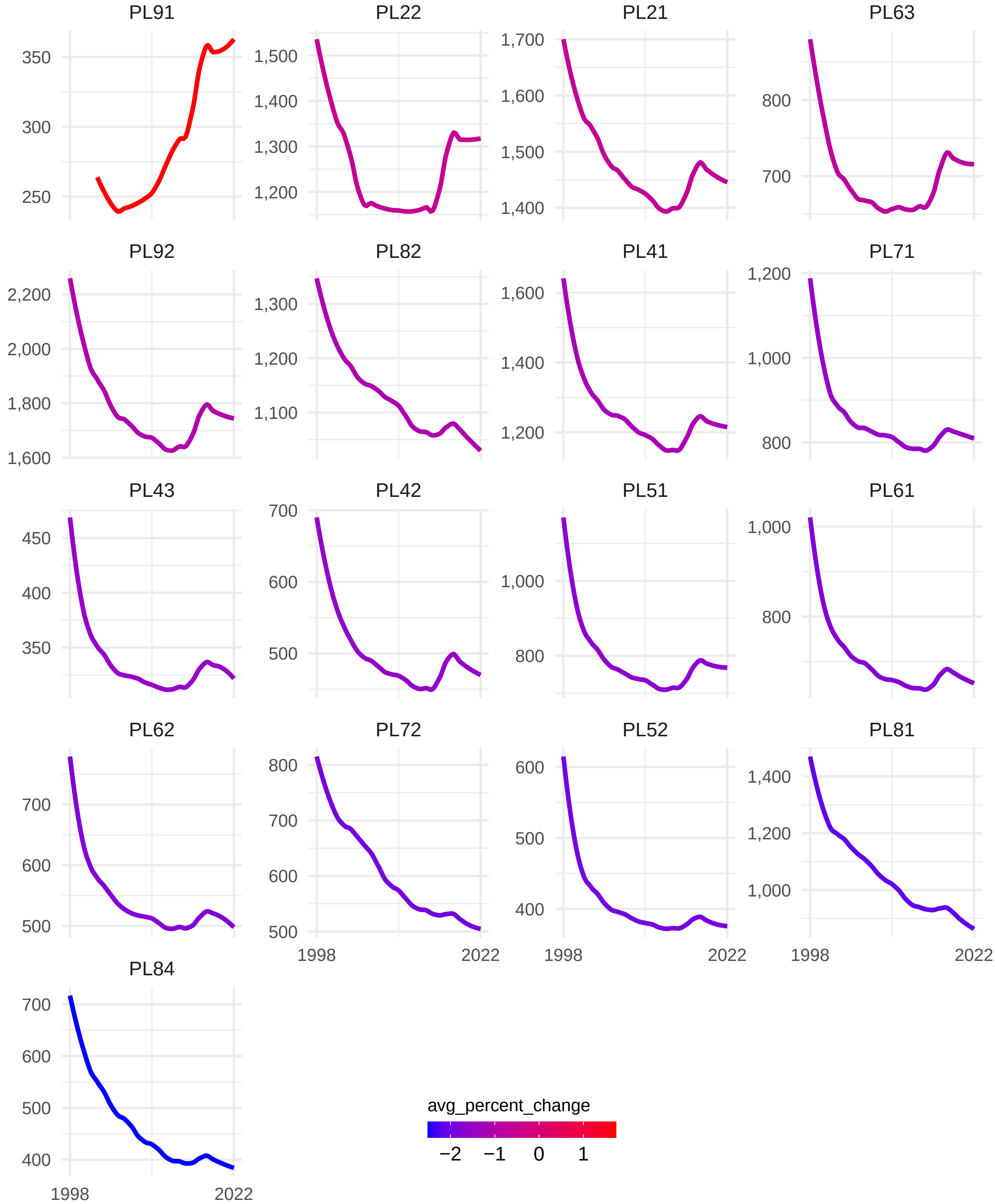
3.2 Number of Institutions

I provide separate overviews of the number of schools, hospitals and third places in Polish NUTS2 regions over time. My thesis assumes that emigration causes these numbers to decrease over time. The following graphs show that in fact only the number of schools decrease and the number of hospitals and third places tend to increase.

3.2.1 Number of Schools in Poland at NUTS2 Level

Apart from Warsaw Metropolitan Area (PL91), all Polish NUTS2 regions see an overall decrease in the number of schools, with some regions displaying an uptick between the years 2015 and 2017. This overall decrease aligns with my assumption that the number of schools decreases over time.

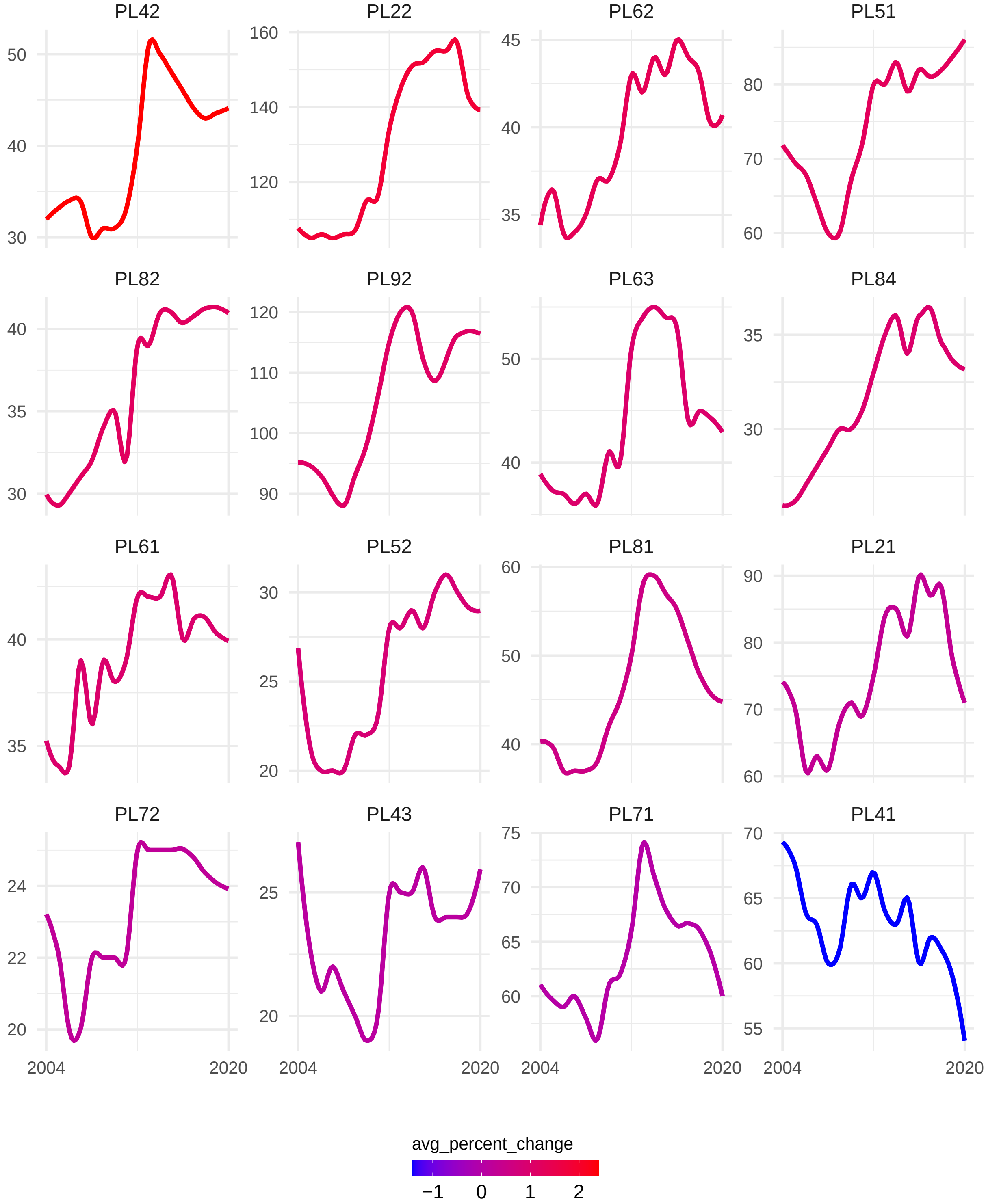
Curves apply smoothing to enhance readability. Plots are ordered by average annual growth rate, with the highest at the top left and decreasing to the right. The growth rate is also indicated by the curve's colour.



3.2.2 Number of Hospitals in Poland at NUTS2 Level

Overall, the number of hospitals increases across Poland, which one exception being Greater Poland Voivodeship (PL41). This goes against my general assumption that the number of hospitals decreases due to emigration.

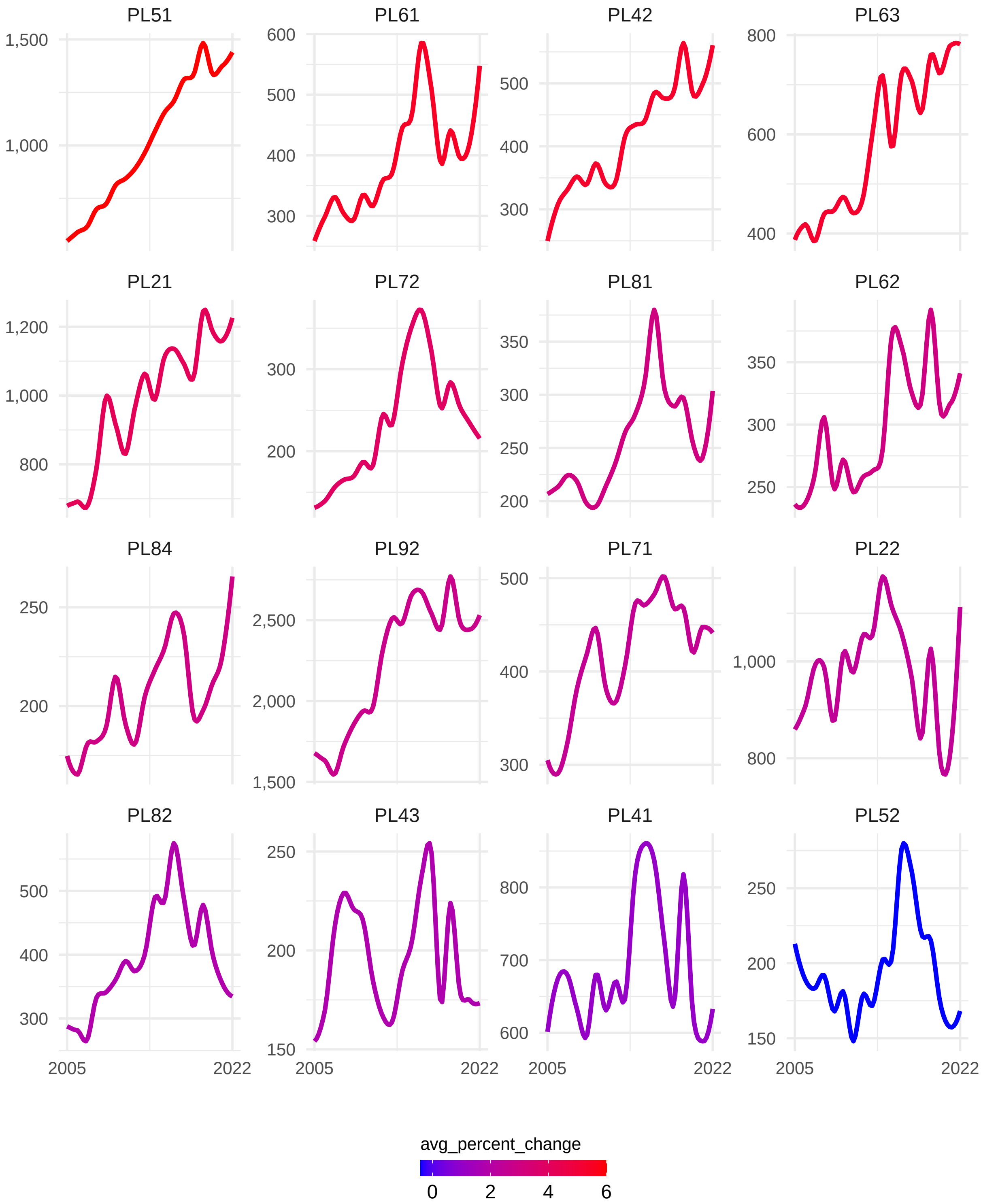
Curves apply smoothing to enhance readability. Plots are ordered by average annual growth rate, with the highest at the top left and decreasing to the right. The growth rate is also indicated by the curve's colour.



3.2.3 Number of Third Places in Poland at NUTS2 Level

There is no consistent trend across all regions when it comes to number of third places in Polish NUTS2 regions. Seen over all regions, the number of third places does not decrease or stays approximately the same.

Curves apply smoothing to enhance readability. Plots are ordered by average annual growth rate, with the highest at the top left and decreasing to the right. The growth rate is also indicated by the curve's colour.



3.3 Ratios of People per Institution

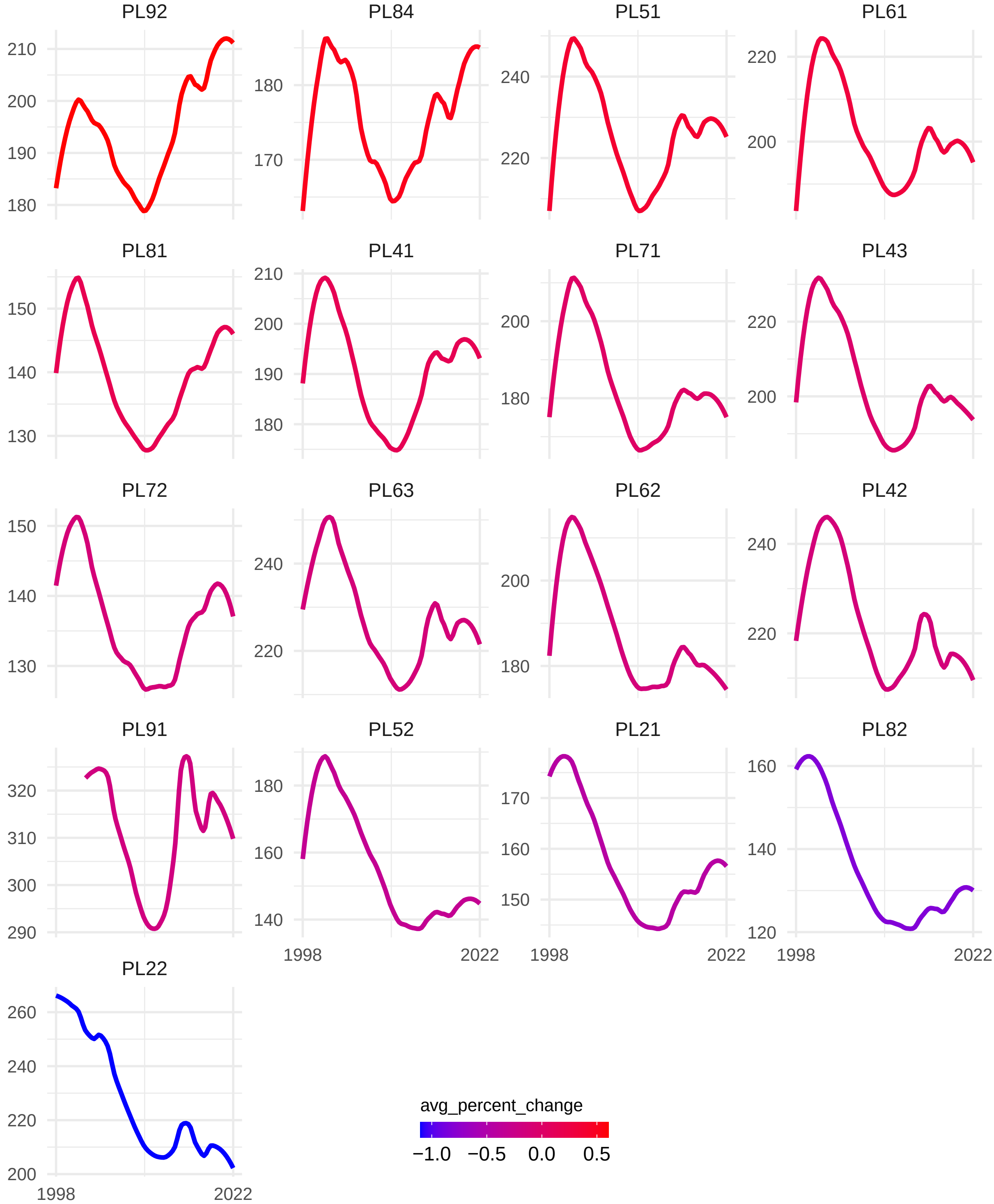
In a next step, I look at the number of people per institution within a region. For example, I count the number of primary school children per school on average, or the number of a region's inhabitants per hospital.

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 inhabitants per hospital.

3.3.1 Number of Primary School Children per School in Poland at NUTS2 Level

The number of primary school children per school sees a dip across regions around the years 2010/2011, after which most regions see an increase. Only the Silesian Voivodeship (PL22) shows a constant decrease.

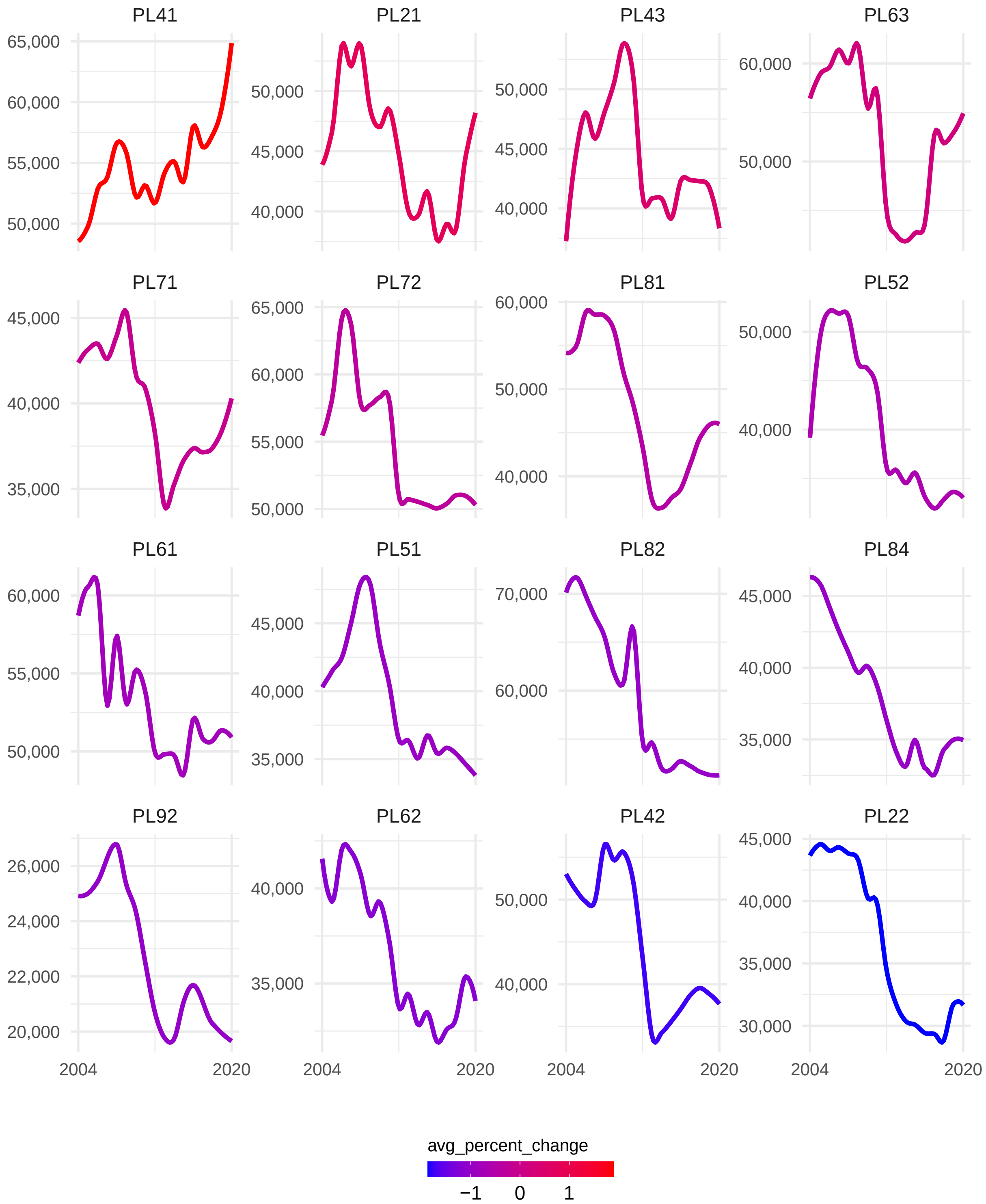
Curves apply smoothing to enhance readability. Plots are ordered by average annual growth rate, with the highest at the top left and decreasing to the right. The growth rate is also indicated by the curve's colour.



3.3.2 Number of People per Hospital in Poland at NUTS2 Level

A general trend across all regions cannot be observed. Only the Greater Poland Voivodeship (PL41) shows a clear increase in number of people per hospital, while all other regions show a decrease.

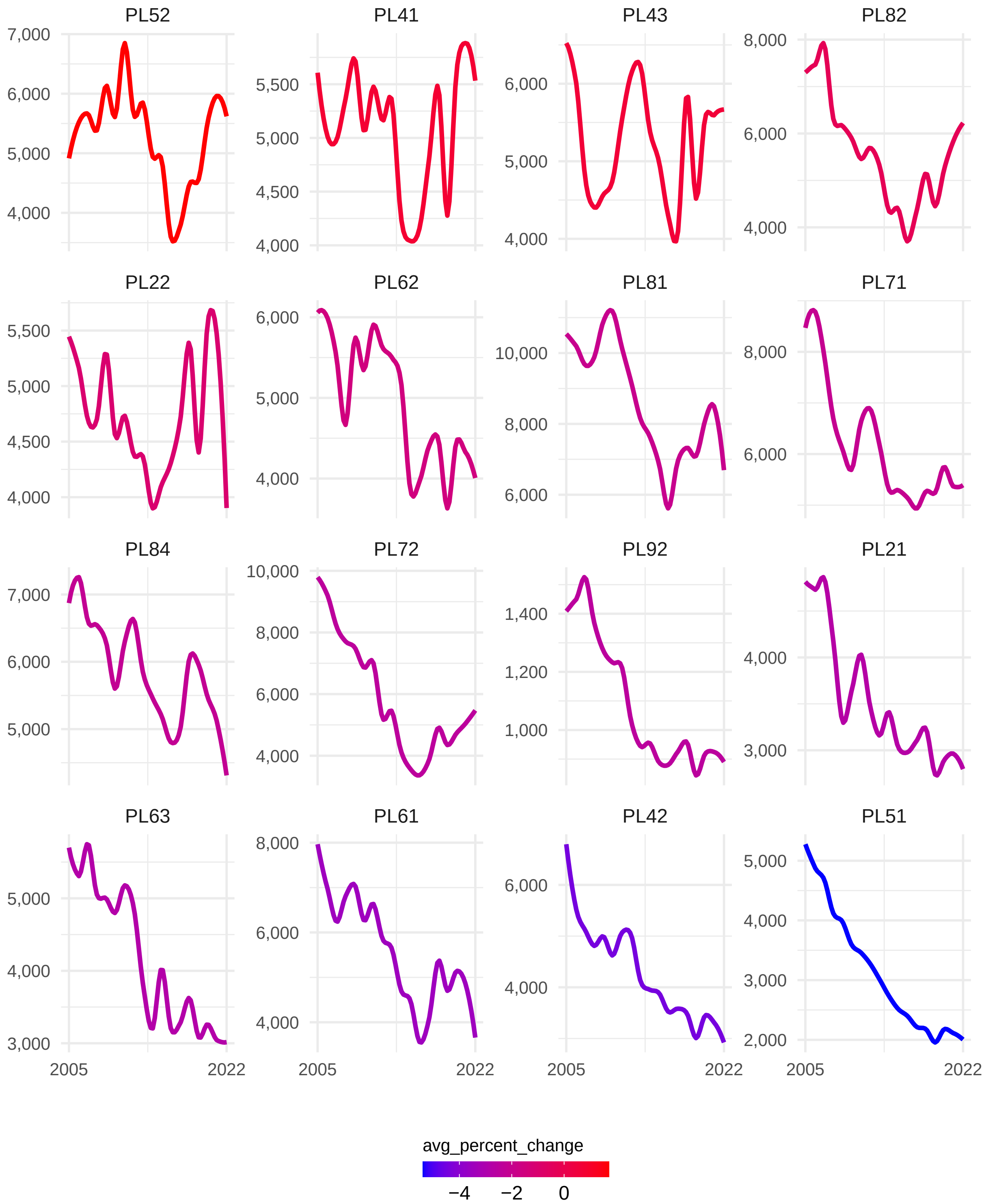
Curves apply smoothing to enhance readability. Plots are ordered by average annual growth rate, with the highest at the top left and decreasing to the right. The growth rate is also indicated by the curve's colour.



3.3.3 Number of People per Third Place in Poland at NUTS2 Level

A clear trend is not visible when taking all regions into account. Generally, when observing the average annual growth rate, we can see that the number of people per third place decreases over time.

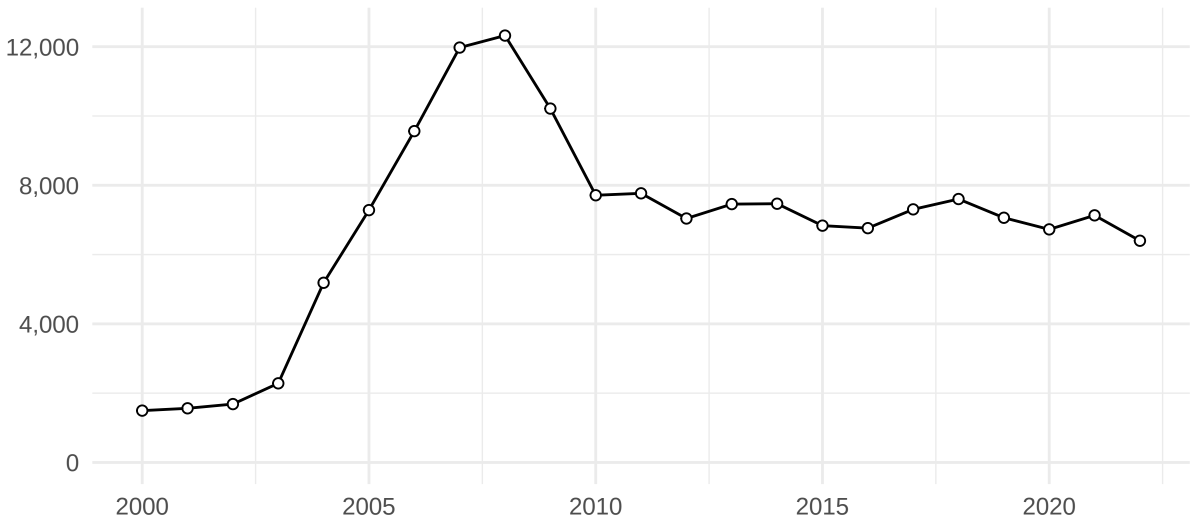
Curves apply smoothing to enhance readability. Plots are ordered by average annual growth rate, with the highest at the top left and decreasing to the right. The growth rate is also indicated by the curve's colour.



3.4 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.

3.4.1 Remittance Inflows (US$ million) for all of Poland



3.4.2 GDP at Current Market Prices in Poland at NUTS2 Level (Euros per Inhabitant)

Each line represents a NUTS2 region in Poland. The Warsaw Metropolitan Area (PL91) displays a consistently higher GDP than all other regions.

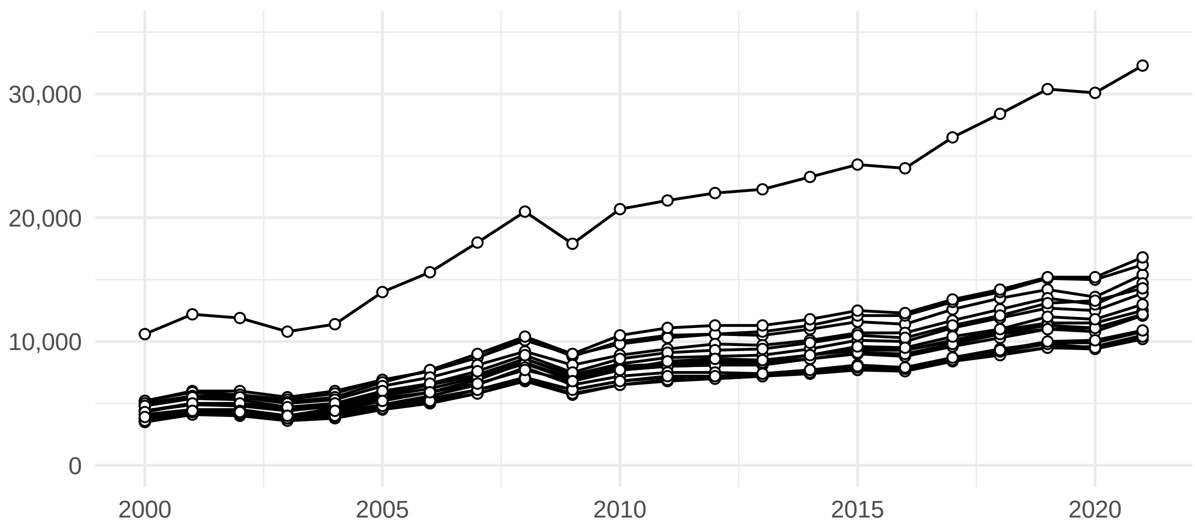


Table 3: DV: Incumbent's Change in Vote Share in Poland

	(1)	(2)	(3)
Children per School	−0.089		
People per Hospital		0.000	
People per Third Place			−0.001*
Avg. Emigration between Elections	0.128	−0.055	−0.152
GDP	0.000	0.000	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

3.5 Modelling Anti-Incumbent Voting in Poland at NUTS2 Level using Ratios

I first model separate ratios on incumbent vote change while controlling for emigration, GDP and remittances. Models 1, 2 and 3 in Table 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 per 1000 inhabitants. Remittances are removed due to collinearity.

Two coefficients point in the right direction but overall one variable is 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 at conventional levels ($p = 0.197$). The ratio of people per hospital displays no effect on incumbent vote share change, while the number of people per third place has a very small yet statistically significant effect.

Table 4: DV: Incumbent's Change in Vote Share in Poland

	(1)	(2)	(3)	(4)
Children per School	−0.109+	−0.092		−0.112*
People per Hospital	0.000		0.000	0.000
People per Third Place		−0.001+	−0.001*	−0.001+
Avg. Emigration between Elections	0.063	−0.038	−0.218	−0.100
GDP	0.000	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 (Table 4, models 1, 2 and 3) and finally combine all ratios in model 4 (see Table 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.

Table 5: DV: Incumbent's Change in Vote Share in Poland

	(1)	(2)	(3)
Number of Schools	0.006		
Number of Hospitals		−0.027	
Number of Third Places			−0.002
Avg. Emigration between Elections	0.028	−0.002	0.035
GDP	−0.001**	0.000	0.000
Num.Obs.	195	186	186
R2	0.502	0.502	0.502
R2 Adj.	0.438	0.438	0.438
R2 Within	0.004	0.000	0.000
R2 Within Adj.	−0.013	−0.018	−0.018
AIC	1504.5	1439.6	1439.6
BIC	1579.8	1510.6	1510.6
RMSE	10.18	10.31	10.31
Std.Errors	by: nuts2016	by: nuts2016	by: nuts2016
FE: nuts2016	X	X	X
FE: year	X	X	X

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

3.6 Modelling Anti-Incumbent Voting in Poland at NUTS2 Level using Number of Institutions

As an alternative to using number of people per institution as the independent variable, I instead use the actual number of institutions. I do this to examine if the closing of an institution has an effect on the incumbent's vote share, regardless of the number of people who then use the institution.

Analogue to Tables 3 and 4, I first model each institution separately (Table 5) and then combine the different independent variables (Table 6). In contrast to using ratios as independent variables, using actual numbers should result in positive coefficients in order to align with my theory. This is because an increase in the number of, e.g., schools, should be associated with an increase in the incumbent's change in vote share – and reversely – a decrease in schools should be associated with a decrease in vote share.

In sum, no model shows statistically significant results and the coefficients signs mostly do not align with my theory.

Table 6: DV: Incumbent's Change in Vote Share in Poland

	(1)	(2)	(3)	(4)
Number of Schools	0.004	0.003		0.003
Number of Hospitals	−0.031		−0.016	−0.020
Number of Third Places		−0.002	−0.001	−0.001
Avg. Emigration between Elections	0.008	0.044	0.022	0.030
GDP	0.000	0.000	0.000	0.000
Num.Obs.	186	186	186	186
R2	0.502	0.502	0.502	0.502
R2 Adj.	0.434	0.434	0.434	0.431
R2 Within	0.000	0.000	0.000	0.000
R2 Within Adj.	−0.024	−0.024	−0.024	−0.031
AIC	1441.6	1441.6	1441.6	1443.6
BIC	1515.8	1515.8	1515.8	1521.0
RMSE	10.31	10.31	10.31	10.31
Std.Errors	by: nuts2016	by: nuts2016	by: nuts2016	by: nuts2016
FE: nuts2016	X	X	X	X
FE: year	X	X	X	X

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

3.7 Modelling with Truncated Data

Do models improve when truncating parts of the data? I truncate the data in three different ways to explore if results vary. Model 1 uses data where observations with a small vote change, between -2 and 2, are removed. This is done to see if extreme vote change values lead to more significant results. Model 2 uses data in which observations with extremely high (over 25%) and extremely low (under 5%) vote shares are removed. Model 3 uses data where high levels of emigration rates (above 4 per 1000 inhabitants) are removed. Both models 2 and 3 attempt to remove outliers in order to build models with higher validity.

Model results using truncated data are shown in Tables 7 and 8. The models in Table 7 use ratios as independent variables, models in Table 8 use numbers of institutions.

Truncated model 1, which focuses on extreme values, shows the most promising results when using ratios as independent variables (see Table 7). The school ratio and third places ratio coefficients become more pronounced and have higher statistical significance compared to the models using all observations. The emigration coefficient also becomes statistically significant, which is an improvement compared to the non-truncated models.

The models that use actual numbers of institutions as independent variables display less promising results (see Table 8). Almost all coefficients are not significant at conventional levels. The one statistically significant value shows a negative coefficient, meaning that an increase in number of third places is associated with a decrease in incumbent vote share, which does not align with my theory.

Table 7: DV: Incumbent's Change in Vote Share in Poland (Truncated Data)

	(1)	(2)	(3)
Children per School	−0.379**	−0.196	−0.078
People per Hospital	0.000+	0.000	0.000
People per Third Place	−0.003*	−0.001	−0.001+
Avg. Emigration between Elections	−0.720*	−1.371**	0.185
GDP	0.001	0.000	−0.001
Num.Obs.	147	70	173
R2	0.725	0.529	0.506
R2 Adj.	0.674	0.309	0.430
R2 Within	0.074	0.013	0.005
R2 Within Adj.	0.037	−0.092	−0.028
AIC	1086.5	554.5	1343.0
BIC	1158.2	606.2	1418.7
RMSE	8.28	9.14	10.22
Std.Errors	by: nuts2016	by: nuts2016	by: nuts2016
FE: nuts2016	X	X	X
FE: year	X	X	X

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

Table 8: DV: Incumbent's Change in Vote Share in Poland (Truncated Data)

	(1)	(2)	(3)
Number of Schools	0.028	0.032	0.005
Number of Hospitals	−0.125	0.155	−0.037
Number of Third Places	−0.003	−0.019+	−0.001
Avg. Emigration between Elections	−0.409	−0.872+	0.764
GDP	0.001	0.003	0.000
Num.Obs.	147	70	173
R2	0.707	0.530	0.504
R2 Adj.	0.652	0.310	0.427
R2 Within	0.013	0.015	0.001
R2 Within Adj.	−0.028	−0.090	−0.033
AIC	1096.0	554.3	1343.8
BIC	1167.8	606.1	1419.5
RMSE	8.55	9.14	10.24
Std.Errors	by: nuts2016	by: nuts2016	by: nuts2016
FE: nuts2016	X	X	X
FE: year	X	X	X

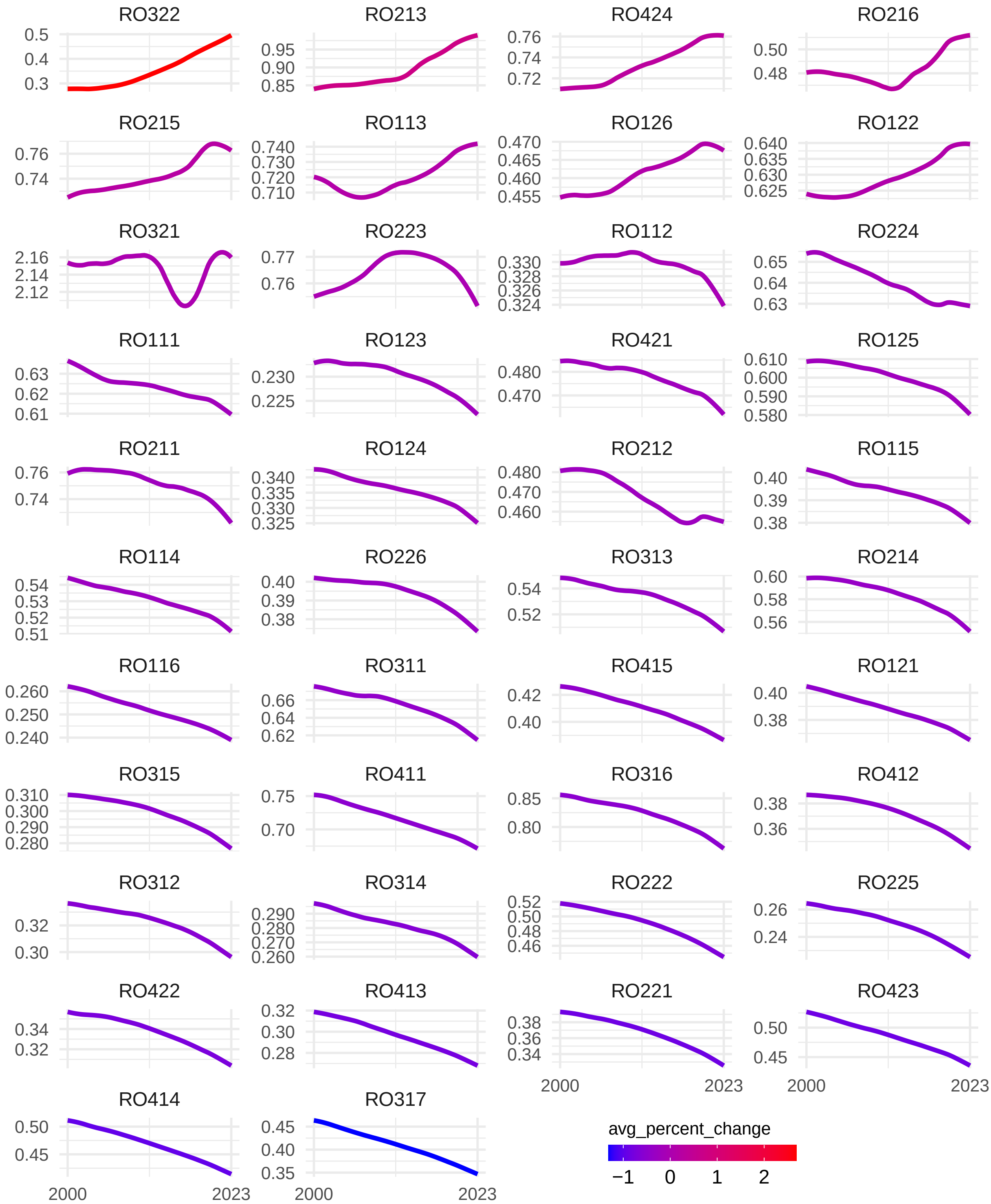
+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

4 Service Cuts in Romania

4.1 Population and Emigration

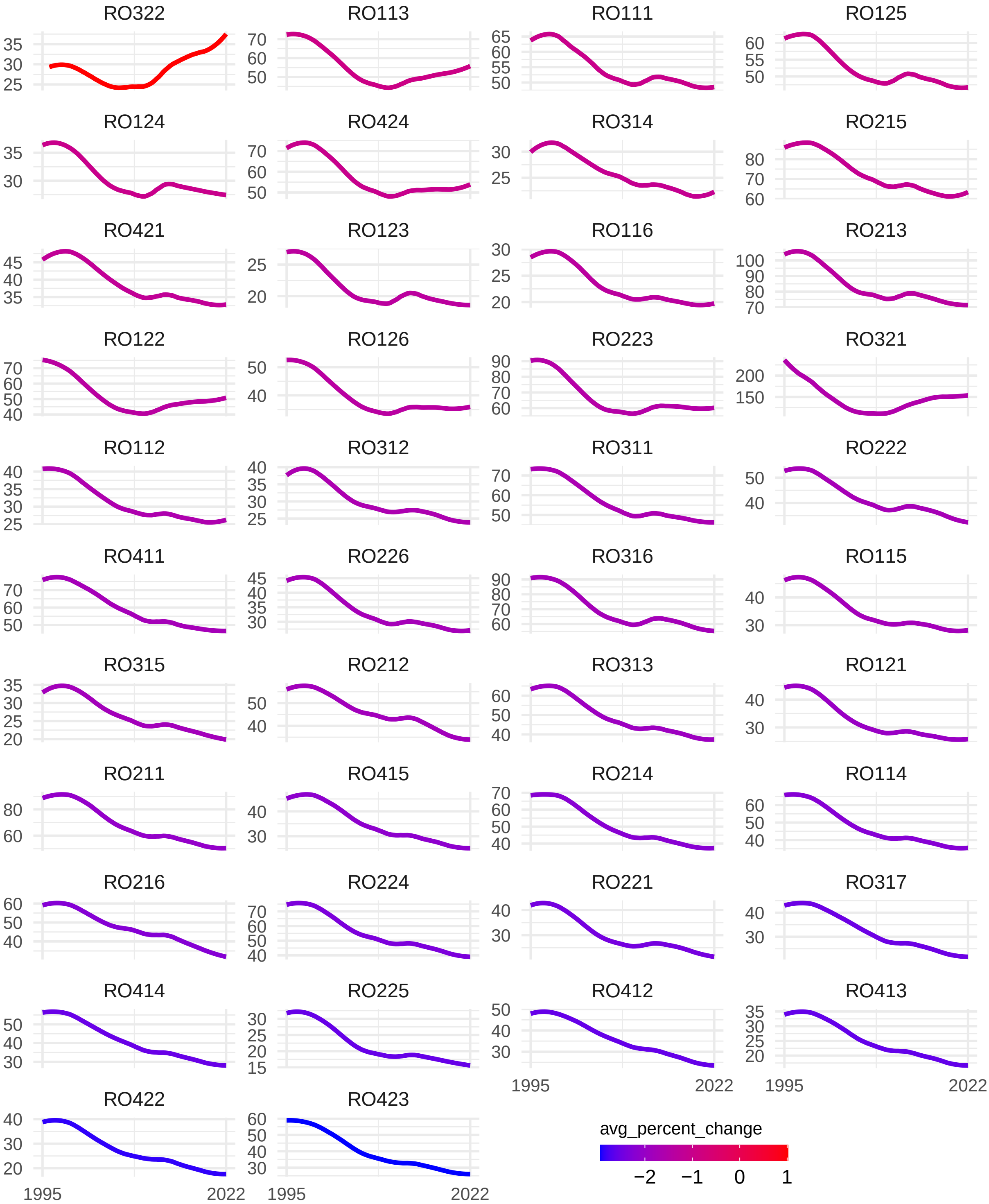
4.1.1 Population in Millions in Romania at NUTS3 Level

Curves apply smoothing to enhance readability. Plots are ordered by average annual growth rate, with the highest at the top left and decreasing to the right. The growth rate is also indicated by the curve's colour.



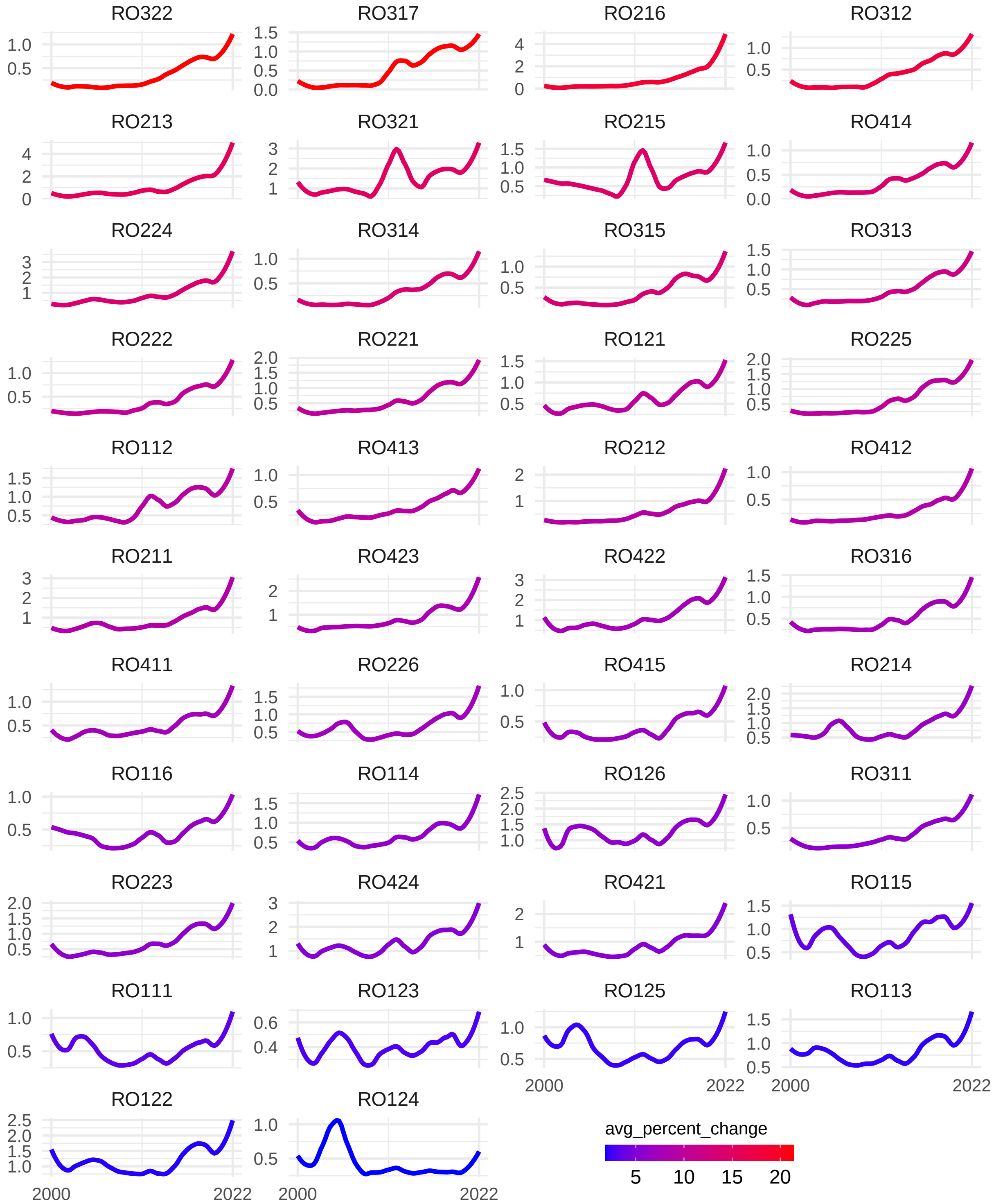
4.1.2 Primary Age Population in 1000s in Romania at NUTS3 Level

Curves apply smoothing to enhance readability. Plots are ordered by average annual growth rate, with the highest at the top left and decreasing to the right. The growth rate is also indicated by the curve's colour.



4.1.3 Emigration per 1000 Inhabitants from Romania at NUTS3 Level

Curves apply smoothing to enhance readability. Plots are ordered by average annual growth rate, with the highest at the top left and decreasing to the right. The growth rate is also indicated by the curve's colour.

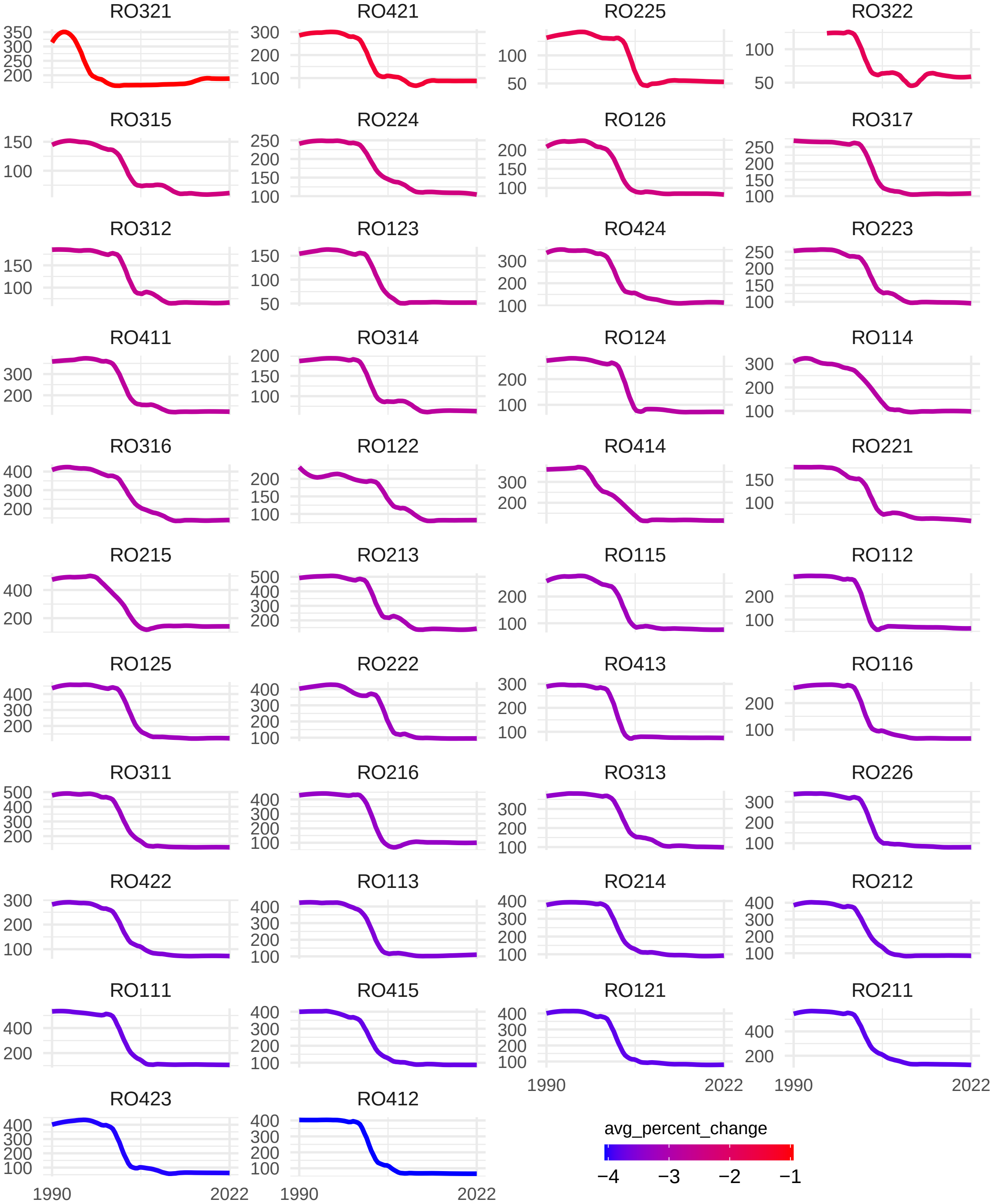


4.2 Number of Institutions

4.2.1 Number of Schools in Romania at NUTS3 Level

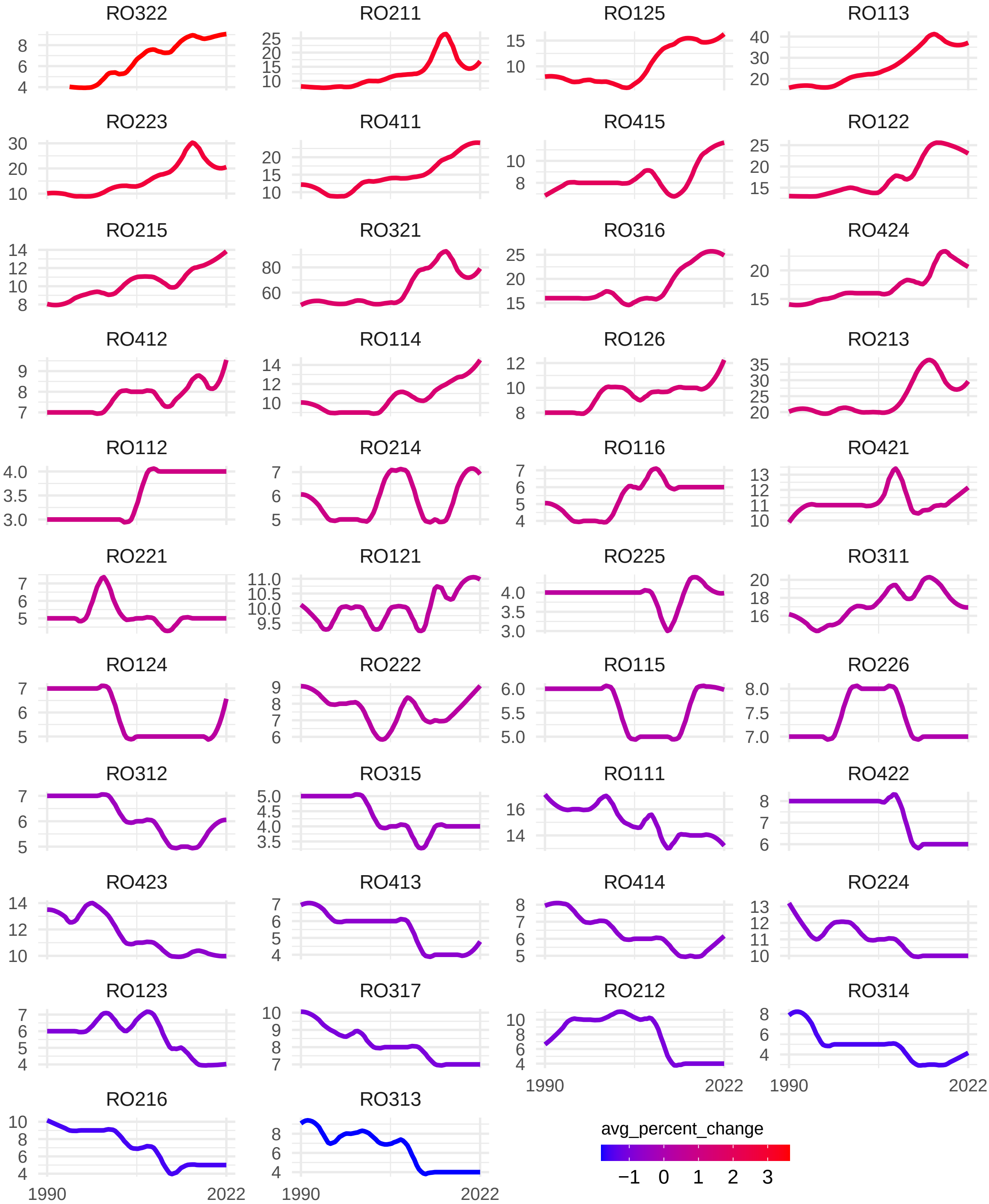
Many regions see a sudden drop in number of schools during the beginning of the 2000s, the reason for which is currently unclear.

Curves apply smoothing to enhance readability. Plots are ordered by average annual growth rate, with the highest at the top left and decreasing to the right. The growth rate is also indicated by the curve's colour.



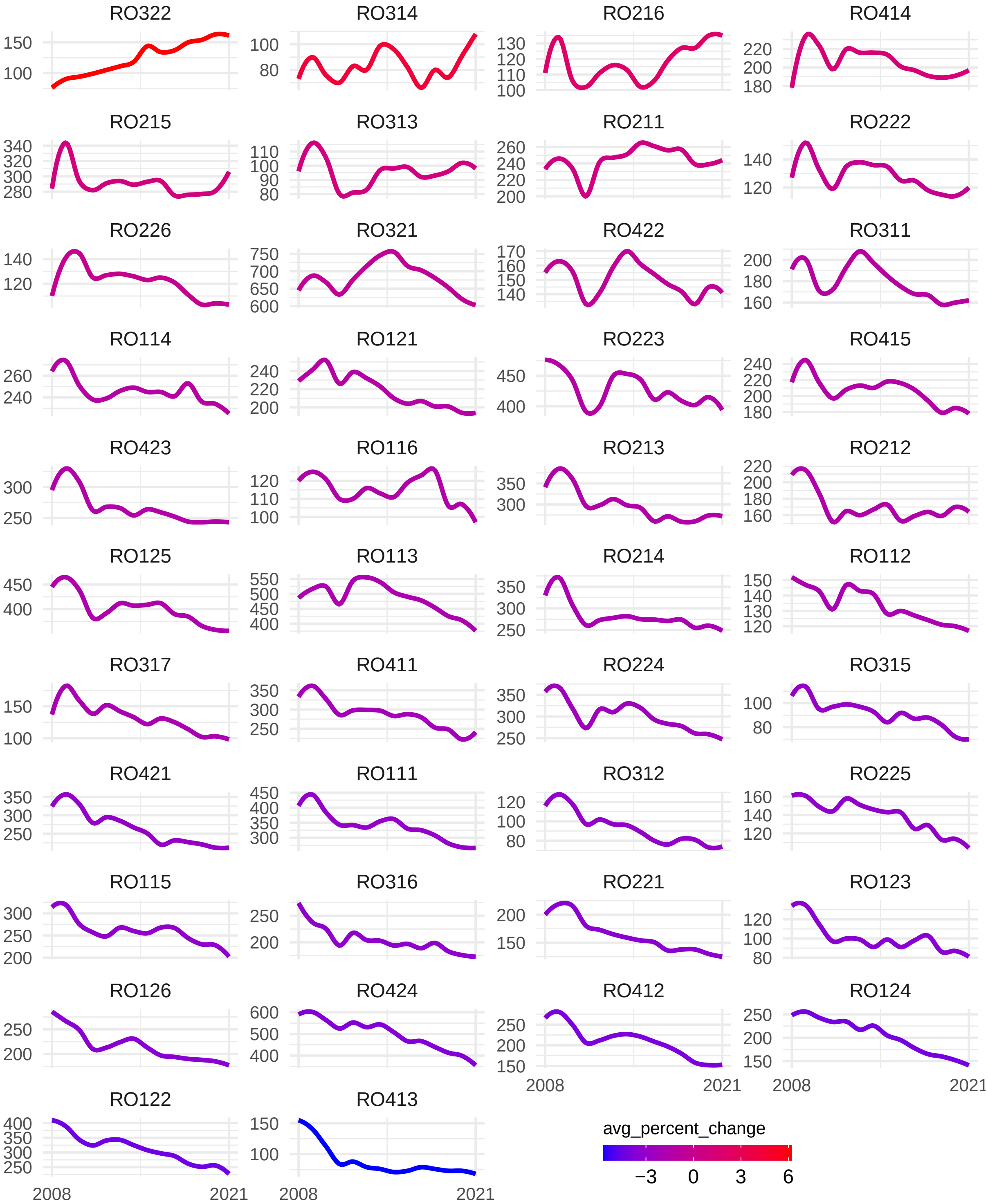
4.2.2 Number of Hospitals in Romania at NUTS3 Level

Curves apply smoothing to enhance readability. Plots are ordered by average annual growth rate, with the highest at the top left and decreasing to the right. The growth rate is also indicated by the curve's colour.



4.2.3 Number of Third Places in Romania at NUTS3 Level

Curves apply smoothing to enhance readability. Plots are ordered by average annual growth rate, with the highest at the top left and decreasing to the right. The growth rate is also indicated by the curve's colour.



4.3 Ratios of People per Institution

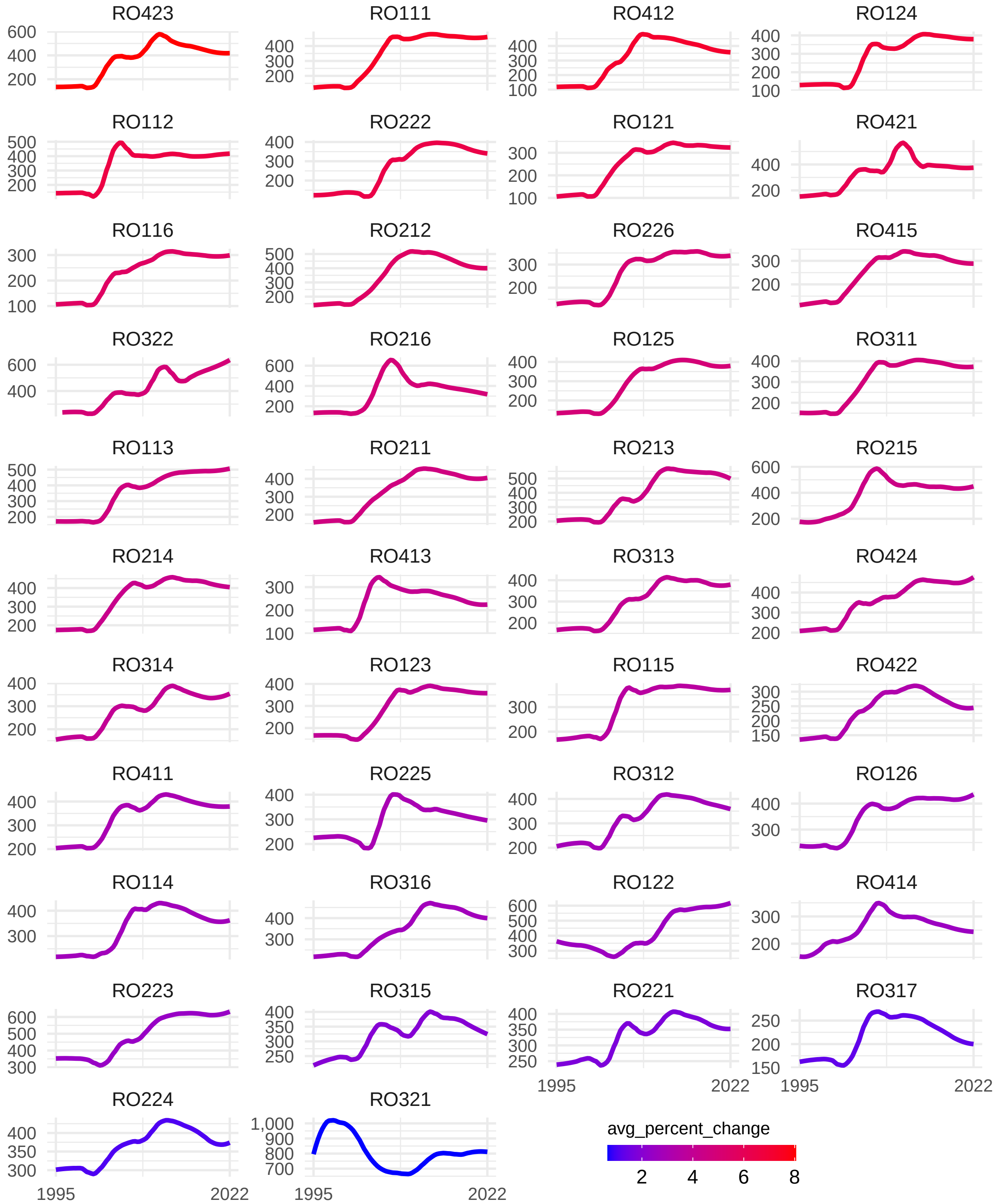
In a next step, I look at the number of people per institution within a region. For example, I count the number of primary school children per school on average, or the number of a region's inhabitants per hospital.

The following plots do not show a consistent trends across all regions in Romania. A possible next step could be to categorise NUTS3 regions, e.g., into urban and rural regions, or regions that border neighbouring countries, in an attempt to find groups with similar trends.

4.3.1 Number of Primary School Children per School in Romania at NUTS3 Level

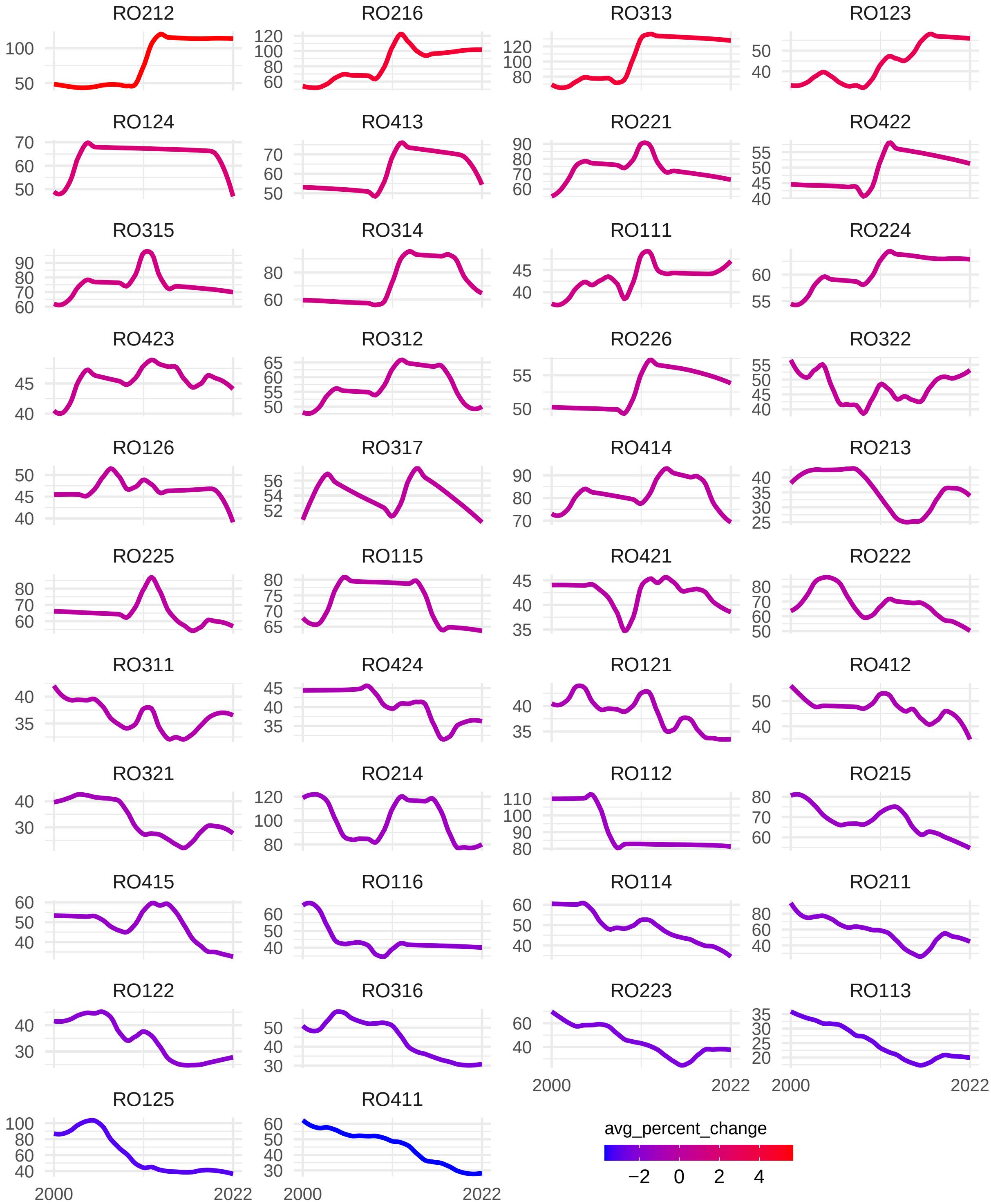
We can generally observe an increase in children per school in Romania. This increase seems to follow the sudden decrease in schools at the beginning of the 2000s.

Curves apply smoothing to enhance readability. Plots are ordered by average annual growth rate, with the highest at the top left and decreasing to the right. The growth rate is also indicated by the curve's colour.



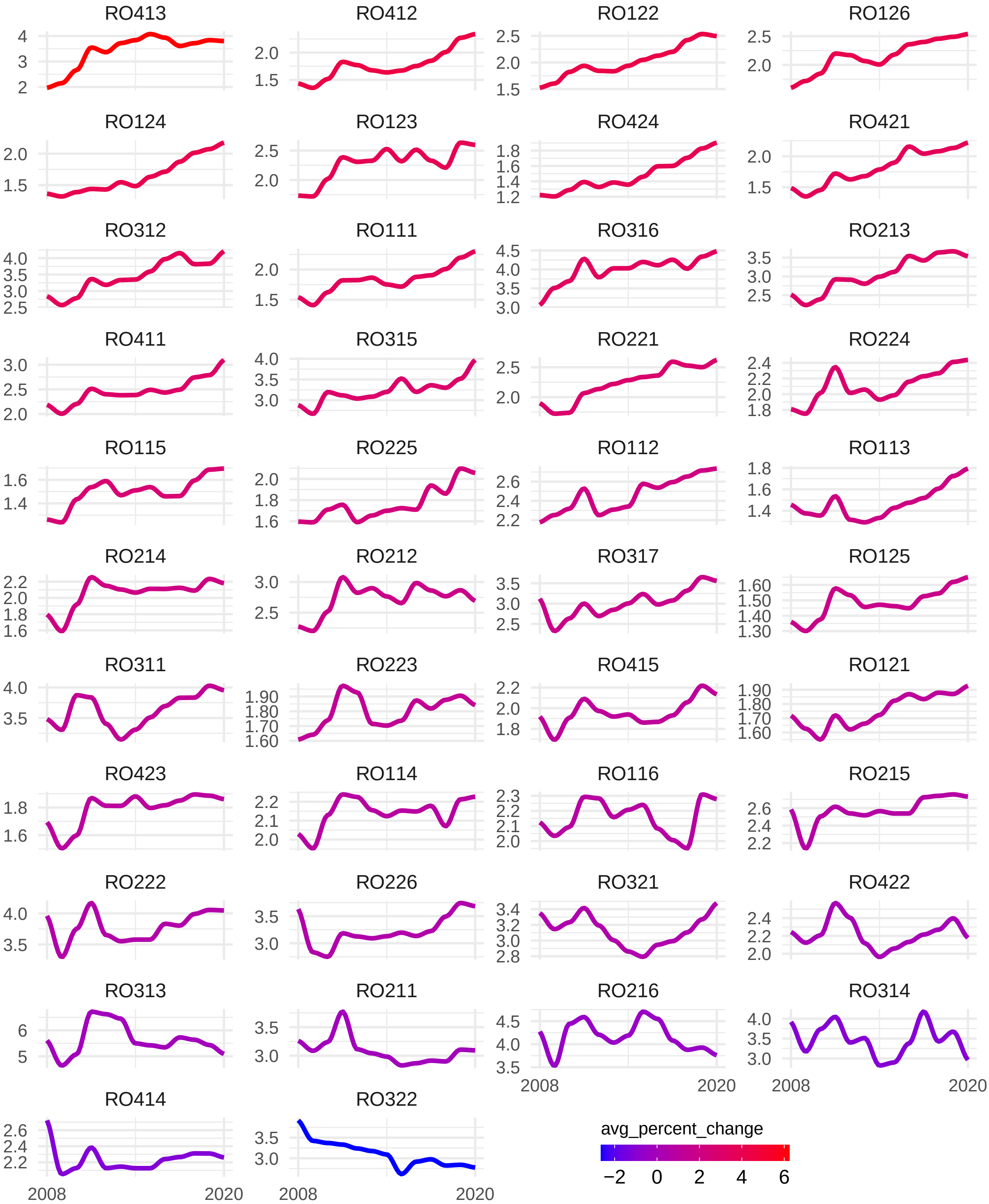
4.3.2 Number of People in 1000s per Hospital in Romania at NUTS3 Level

Curves apply smoothing to enhance readability. Plots are ordered by average annual growth rate, with the highest at the top left and decreasing to the right. The growth rate is also indicated by the curve's colour.



4.3.3 Number of People in 1000s per Third Place in Romania at NUTS3 Level

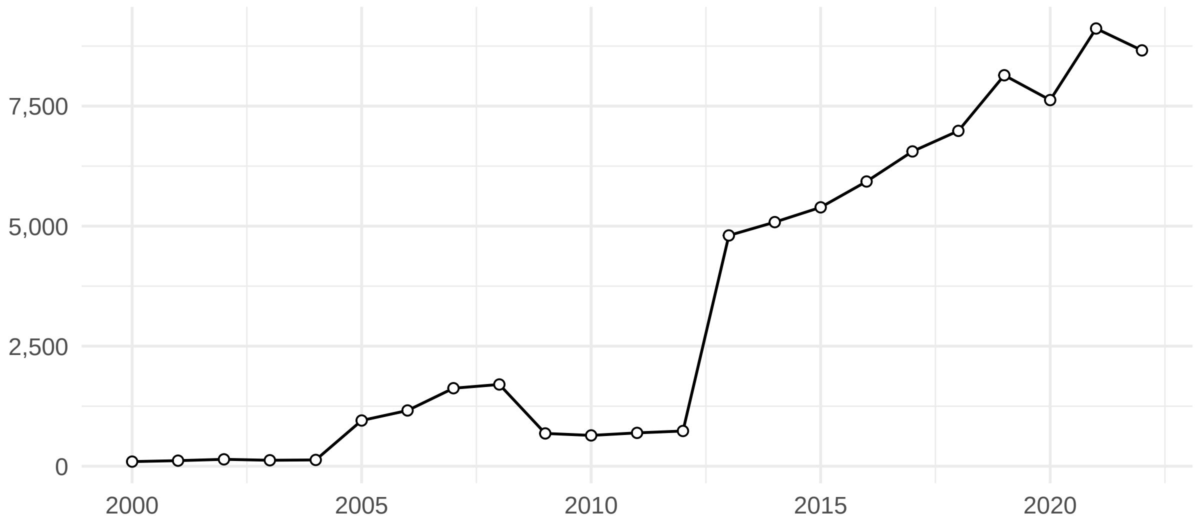
Curves apply smoothing to enhance readability. Plots are ordered by average annual growth rate, with the highest at the top left and decreasing to the right. The growth rate is also indicated by the curve's colour.



4.4 Control Variables

I use these data as control variables in my models. Data on remittances are only available at the national level.

4.4.1 Remittance Inflows (US$ million) for all of Romania



4.4.2 GDP at Current Market Prices in Romania at NUTS3 Level (Euros per Inhabitant)

Each line represents a NUTS3 region in Romania. Bucharest (RO321) displays a consistently higher GDP than all other regions.

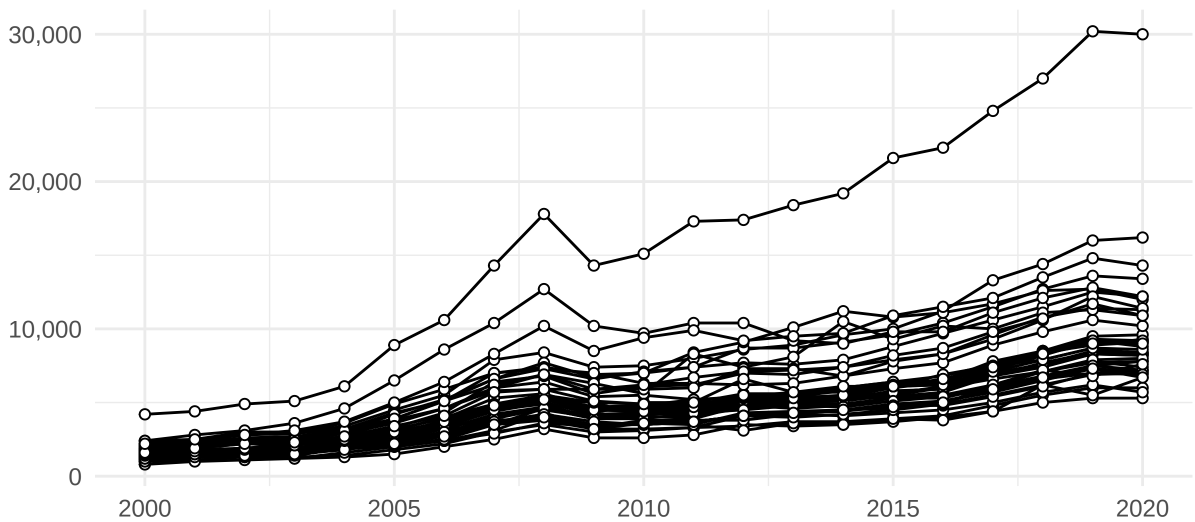


Table 9: DV: Incumbent's Change in Vote Share in Romania

	(1)	(2)	(3)
Children per School	−0.021		
People per Hospital		0.000	
People per Third Places			0.002
Avg. Emigration between Elections	−7.136***	−6.718***	−6.690***
GDP	0.004*	0.003*	0.003*
Num.Obs.	168	168	168
R2	0.803	0.802	0.802
R2 Adj.	0.730	0.729	0.729
R2 Within	0.186	0.183	0.183
R2 Within Adj.	0.166	0.163	0.163
AIC	1239.6	1240.3	1240.2
BIC	1383.4	1384.0	1383.9
RMSE	7.36	7.38	7.38
Std.Errors	by: nuts2016	by: nuts2016	by: nuts2016
FE: nuts2016	X	X	X
FE: year	X	X	X

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

4.5 Modelling Anti-Incumbent Voting in Romania at NUTS3 Level using Ratios

I first model separate ratios on incumbent vote change while controlling for emigration, GDP and remittances. Models 1, 2 and 3 in Table 9 each use a ratio separately. NUTS3 regions and years are used as fixed effects in all models. The emigration rate is the rolling emigration average between two national elections per 1000 inhabitants. Remittances are removed due to collinearity.

Only model 1, which uses the ratio of children per school, has a negative coefficient, that indicates that more children per school is associated with a loss in incumbent vote share. However, this coefficient is not statistically significant at conventional levels ($p = 0.26$). The ratio of people per hospital has no effect on incumbent vote share and the number of people per third place shows a positive coefficient.

A highlight is the emigration control variable, which displays a relatively strong and statistically significant effect. Higher levels of emigration between elections is associated with a decrease in incumbent vote share across all three models.

Table 10: DV: Incumbent's Change in Vote Share in Romania

	(1)	(2)	(3)	(4)
Children per School	−0.026	−0.020		−0.026
People per Hospital	0.000		0.000	0.000
People per Third Places		0.002	0.003	0.003
Avg. Emigration between Elections	−6.945***	−6.933***	−6.433***	−6.663***
GDP	0.003+	0.004*	0.003+	0.003+
Num.Obs.	168	168	168	168
R2	0.804	0.804	0.803	0.806
R2 Adj.	0.730	0.729	0.729	0.730
R2 Within	0.193	0.190	0.189	0.199
R2 Within Adj.	0.167	0.163	0.162	0.166
AIC	1240.2	1240.9	1241.1	1241.0
BIC	1387.0	1387.7	1387.9	1390.9
RMSE	7.33	7.35	7.35	7.31
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 (Table 10, models 1, 2 and 3) and finally combine all ratios in model 4 (see Table 10). Again, NUTS3 regions and years are used as fixed effects.

Results are similar to models 1 through 3 in Table 9: The coefficients of the independent variables are not statistically significant at conventional levels and, apart from children per school, are positive and thus do not align with my theory. The ratio of children per school in model 4 becomes statistically more significant compared to model 1 in Table 9, but at $p = 0.177$ is still above the 10% threshold for conventional statistical significance.

As with the models in Table 9, we can observe that the emigration control variable in Table 10 again has a strong and statistically significant effect across all models.

Table 11: DV: Incumbent's Change in Vote Share in Romania

	(1)	(2)	(3)
Number of Schools	0.213*		
Number of Hospitals		−0.095	
Number of Third Places			−0.022
Avg. Emigration between Elections	−7.523***	−6.773***	−6.502***
GDP	0.004**	0.004+	0.003*
Num.Obs.	168	168	168
R2	0.812	0.801	0.801
R2 Adj.	0.742	0.728	0.728
R2 Within	0.224	0.180	0.181
R2 Within Adj.	0.205	0.160	0.161
AIC	1231.7	1240.9	1240.7
BIC	1375.4	1384.6	1384.4
RMSE	7.19	7.39	7.39
Std.Errors	by: nuts2016	by: nuts2016	by: nuts2016
FE: nuts2016	X	X	X
FE: year	X	X	X

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

4.6 Modelling Anti-Incumbent Voting in Romania at NUTS3 Level using Number of Institutions

As an alternative to using number of people per institution as the independent variable, I instead use the actual number of institutions. I do this to examine if the closing of an institution has an effect on the incumbent's vote share, regardless of the number of people who then use the institution.

Analogue to Tables 9 and 10, I first model each institution separately (Table 11) and then combine the different independent variables (Table 12). In contrast to using ratios as independent variables, using actual numbers should result in positive coefficients in order to align with my theory. This is because an increase in the number of, e.g., schools, should be associated with an increase in the incumbent's change in vote share – and reversely – a decrease in schools is associated with a decrease in vote share.

Using the number of schools as an independent variable is statistically significant across all models and always displays a positive coefficient (see Tables 11 and 12). This means that an decrease in the number of schools is associated with a decrease in the incumbent's vote share. In short, less schools in a region is bad for incumbent parties. All models also display relatively high R2 values, meaning that all models can explain a large proportion of the dependent variable's variance.

Table 12: DV: Incumbent's Change in Vote Share in Romania

	(1)	(2)	(3)	(4)
Number of Schools	0.216*	0.240*		0.263**
Number of Hospitals	0.037		−0.052	0.223
Number of Third Places		−0.053	−0.019	−0.069
Avg. Emigration between Elections	−7.578***	−6.665***	−6.489***	−6.734***
GDP	0.004*	0.003*	0.003+	0.002
Num.Obs.	168	168	168	168
R2	0.812	0.814	0.801	0.815
R2 Adj.	0.740	0.744	0.726	0.743
R2 Within	0.224	0.234	0.181	0.238
R2 Within Adj.	0.198	0.209	0.154	0.206
AIC	1233.7	1231.5	1242.7	1232.6
BIC	1380.5	1378.3	1389.5	1382.6
RMSE	7.19	7.15	7.39	7.13
Std.Errors	by: nuts2016	by: nuts2016	by: nuts2016	by: nuts2016
FE: nuts2016	X	X	X	X
FE: year	X	X	X	X

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

4.7 Modelling with Truncated Data

Do models improve when truncating parts of the data? I truncate the data in three different ways to explore if results vary. Model 1 uses data where observations with a small vote change, between -0.5 and 0.5, are removed. This is done to see if extreme vote change values lead to more significant results. Model 2 uses data in which observations with extremely high (over 30%) and extremely low (under 5%) vote shares are removed. Model 3 uses data where high levels of emigration rates (above 2 per 1000 inhabitants) are removed. Both models 2 and 3 attempt to remove outliers in order to build models with higher validity.

Model results using truncated data are shown in Tables 13 and 14. The models in Table 13 use ratios as independent variables, models in Table 14 use numbers of institutions.

None of the truncated models show particularly promising results, apart model 1 when using actual numbers of institutions (see Table 14). As with the non-truncated data, the number of schools and incumbent change in vote share are positively associated, with a decrease in schools going hand in hand with a decrease in incumbent vote share. All other coefficients from all other models using truncated data do not display statistically significant results. Also as with the non-truncated data, the control variable for emigration displays highly significant and strong effects, with an increase in emigration being associated with a decrease in incumbent vote share.

Table 13: DV: Incumbent's Change in Vote Share in Romania (Truncated Data)

	(1)	(2)	(3)
Children per School	−0.028	−0.007	0.059
People per Hospital	0.000	0.000	0.000
People per Third Places	0.001	−0.004	0.003
Avg. Emigration between Elections	−7.421***	−3.012*	−17.532*
GDP	0.004+	0.004**	0.001
Num.Obs.	117	84	94
R2	0.823	0.904	0.846
R2 Adj.	0.703	0.790	0.758
R2 Within	0.237	0.214	0.181
R2 Within Adj.	0.181	0.110	0.111
AIC	899.4	613.9	684.3
BIC	1031.9	725.7	773.3
RMSE	7.50	5.41	6.35
Std.Errors	by: nuts2016	by: nuts2016	by: nuts2016
FE: nuts2016	X	X	X
FE: year	X	X	X

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

Table 14: DV: Incumbent's Change in Vote Share in Romania (Truncated Data)

	(1)	(2)	(3)
Number of Schools	0.235*	0.170	0.150
Number of Hospitals	0.131	0.408	0.001
Number of Third Places	−0.070	−0.004	0.019
Avg. Emigration between Elections	−7.166***	−4.329*	−14.743+
GDP	0.003	0.003	0.002
Num.Obs.	117	84	94
R2	0.832	0.908	0.843
R2 Adj.	0.717	0.800	0.752
R2 Within	0.274	0.250	0.162
R2 Within Adj.	0.221	0.151	0.091
AIC	893.5	609.9	686.4
BIC	1026.1	721.7	775.4
RMSE	7.31	5.28	6.42
Std.Errors	by: nuts2016	by: nuts2016	by: nuts2016
FE: nuts2016	X	X	X
FE: year	X	X	X

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

5 Sources

V–Party Dataset

Staffan I. Lindberg, Nils Du ̈pont, Masaaki Higashijima, Yaman Berker Kavasoglu, Kyle L. Marquardt, Michael Bernhard, Holger D ̈oring, Allen Hicken, Melis Laebens, Juraj Medzihorsky, Anja Neundorf, Ora John Reuter, Saskia Ruth–Lovell, Keith R. Weghorst, Nina Wiesehomeier, Joseph Wright, Nazifa Alizada, Paul Bederke, Lisa Gastaldi, Sandra Grahn, Garry Hindle, Nina Ilchenko, Johannes von R ̈omer, Steven Wilson, Daniel Pemstein, Brigitte Seim. 2022. Varieties of Party Identity and Organization (V–Party) Dataset V2. Varieties of Democracy (V–Dem) Project. https://doi.org/10.23696/vpartydsv2 and: Pemstein, Daniel, Kyle L. Marquardt, Eitan Tzelgov, Yi–ting Wang, Juraj Medzihorsky, Joshua Krusell, Farhad Miri, and Johannes von R ̈omer. 2020. "The V-Dem Measurement Model: Latent Variable Analysis for Cross-National and Cross–Temporal Expert-Coded Data". V-Dem Working Paper No. 21. 5th edition. University of Gothenburg: Varieties of Democracy Institute.

EU-NED European NUTS-Level Election Dataset

Schraff, Dominik; Vergioglou, Ioannis; Demirci, Buket Buse, 2022, "EU-NED: The European NUTS-Level Election Dataset", https://doi.org/10.7910/DVN/IQRYP5, Harvard Dataverse, V1, UNF:6:33YFLJUZN1V4UrerfOsbxg== [fileUNF]

Party Facts

https://partyfacts.herokuapp.com/data/partycodes/%5Bpf_party_id%5D/

Bulgaria

https://www.nsi.bg/en/content/3060/migration-population-districts-municipalities-and-sex (retrieved 2 November 2023)

https://www.nsi.bg/en/content/2975/population-districts-municipalities-place-residence-and-sex (retrieved 2 November 2023)

Estonia

RVR01; https://andmed.stat.ee/en/stat/Lepetatud_tabelid__Rahvastik.Arhiiv__Rahvastikus%C3%BCndmused.%20Arhiiv/RVR01 (retrieved 7 November 2023)

RVR02; https://andmed.stat.ee/en/stat/rahvastik__rahvastikusundmused__ranne/RVR02/table/tableViewLayout2 (retrieved 7 November 2023)

RV022: POPULATION, 1 JANUARY by Year, County, Sex and Age group; https://andmed.stat.ee/en/stat/Lepetatud_tabelid__Rahvastik.Arhiiv__Rahvastikun%C3%A4itajad%20ja%20koosseis.%20Arhiiv/RV022/table/tableViewLayout2 (retrieved 7 November 2023)

RV022U: POPULATION, 1 JANUARY. ADMINISTRATIVE DIVISION AS AT 01.01.2018 by Year, County, Sex and Age group; https://andmed.stat.ee/en/stat/rahvastik__rahvastikunaitajad-ja-koosseis__rahvaarv-ja-rahvastiku-koosseis/RV022U/table/tableViewLayout2 (retrieved 7 November 2023)

Croatia

https://podaci.dzs.hr/media/ueajlqe5/stanovnistvo-pregled-po-zupanijama.xlsx (retrieved 3 November 2023)

https://podaci.dzs.hr/media/ueajlqe5/stanovnistvo-pregled-po-zupanijama.xlsx (retrieved 3 November 2023)

Hungary

https://statinfo.ksh.hu/Statinfo/haViewer.jsp (retrieved 7 November 2023)

https://www.ksh.hu/stadat_files/nep/en/nep0034.html (retrieved 16 November 2023)

Lithuania

https://osp.stat.gov.lt/statistiniu-rodikliu-analize#/ (retrieved 9 November 2023)

Latvia

https://data.stat.gov.lv/pxweb/en/OSP_PUB/START__POP__IB__IBE/IBE080/table/tableViewLayout1/

(retrieved 9 November 2023)
https://data.stat.gov.lv/pxweb/en/OSP_PUB/START__POP__IR__IRS/IRS030/table/tableViewLayout1/ (retrieved 9 November 2023)

Poland

Entire Population, Schools, Hospitals, Third Places
https://bdl.stat.gov.pl/bdl/dane/podgrup/tablica (retrieved November 2023)

Primary Age Population
https://ec.europa.eu/eurostat/databrowser/view/demo_r_d2jan__custom_8622439/default/table?lang=en (retrieved 21 November 2023)

Remittances
https://www.knomad.org/data/remittances (retrieved 27 November 2023)

GDP
https://ec.europa.eu/eurostat/databrowser/view/nama_10r_2gdp__custom_8716685/default/table?lang=en (retrieved 27 November 2023)

Romania

All Data
http://statistici.insse.ro:8077/tempo-online/#/pages/tables/insse-table (retrieved November/December 2023)

Slovenia

https://pxweb.stat.si/SiStatData/pxweb/en/Data/Data/05N1042S.px/table/tableViewLayout2/ (retrieved 10 November 2023)
https://pxweb.stat.si/SiStatData/pxweb/en/Data/Data/05E1016S.px/table/tableViewLayout2/ (retrieved 10 November 2023)

Slovakia

https://datacube.statistics.sk/#!/view/en/vbd_dem/om7021rr/v_om7021rr_00_00_00_en (retrieved 10 November 2023)

1999-2019 Chapel Hill Expert Survey (CHES) trend file

Jolly, Seth, Ryan Bakker, Liesbet Hooghe, Gary Marks, Jonathan Polk, Jan Rovny, Marco Steenbergen, and Milada Anna Vachudova. 2022. "Chapel Hill Expert Survey Trend File, 1999-2019." Electoral Studies 75 (February). https://doi.org/10.1016/j.electstud.2021.102420