Models

Project: Dominant Party Regimes in Sub-Saharan Africa

Stepan Polikanov

This file reports the modelling process and the associated analytic decisions. It also includes the results of sensitivity tests.

Read in the data compiled in the *data_building.qmd* file.

```
data <- read_rds(here("data", "data_built", "data.built.rds"))
```

The main and only model in my original work was a multinomial logistic model. The main disadvantage of using this class of models in this context is that logistic models typically require more observations than linear models. Nevertheless, I start with a multinomial model to i) establish a baseline, ii) provide ground to validate the analytical coding of the dominant party regimes and iii) because there are not many alternative methods to analyze the data.

I first use a preferred specification, with party system institutionalization and education variables imputed with the mean of 2013-2018 values for the country with missing data, ELF imputed with previous values for the country and oil rents variable presented as a dummy.

```
m1a <- multinom(party_reg ~ v2xps_party_mean + elf_fill + col_type
                + housesys + v2x_polyarchy + oil_rent_dummy + e_peaveduc_mean,
                data = data)

m1b <- multinom(relevel(party_reg, ref = "Autocratic Dominant")
                 ~ v2xps_party_mean + elf_fill + col_type
                 + housesys + v2x_polyarchy + oil_rent_dummy + e_peaveduc_mean,
                data = data)
```

```
modelsummary(list("Ref: Non-dominant" = m1a,
                  "Ref: Autocratic Dominant" = m1b), stars = T,
             shape = term ~ response, coef_rename = T,
             output = "kableExtra") |>
  kable_styling(latex_options = c("scale_down"))
```

Table 1: Multinomial logistic model, preferred specification

	Ref: Non-dominant		Ref: Autocratic Dominant	
	Democratic Dominant	Autocratic Dominant	Democratic Dominant	Not Dominant
(Intercept)	2.131	8.219	−6.090	−8.239
	(4.319)	(5.973)	(4.960)	(5.977)
PSI (Mean)	11.261+	8.545	2.708	−8.556
	(5.836)	(7.040)	(5.826)	(7.041)
ELF (Imputed)	3.852	3.559	0.305	−3.544
	(3.218)	(3.872)	(3.228)	(3.871)
Colonial Legacy [France]	0.477	−1.215	1.693	1.220
	(1.585)	(1.967)	(1.680)	(1.968)
Colonial Legacy [Mixed]	−7.858	−3.529	−3.489	6.267
	(123.274)	(113.129)	(129.126)	(166.662)
Colonial Legacy [None]	−5.960	2.763	−7.018	−2.748
	(135.665)	(16.287)	(57.399)	(16.298)
Colonial Legacy [Other]	−3.770	−4.459	0.696	4.475
	(3.034)	(4.202)	(3.202)	(4.205)
Electoral Rules [Plural]	−4.987*	−3.997	−0.988	4.005
	(2.525)	(3.831)	(3.121)	(3.830)
Democracy	−15.564*	−26.599**	11.022	26.608**
	(6.615)	(9.945)	(8.265)	(9.945)
Oil Dummy [Not equal to zero]	−0.961	0.111	−1.078	−0.118
	(1.494)	(2.049)	(1.746)	(2.050)
Education (Mean Imputed)	0.164	−0.082	0.246	0.082
	(0.264)	(0.392)	(0.361)	(0.392)
Num.Obs.	38		38	
R2	0.511		0.511	
R2 Adj.	0.488		0.488	
AIC	86.5		86.5	
BIC	122.5		122.5	
RMSE	0.34		0.34	

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

```
check_model(m1a)
```

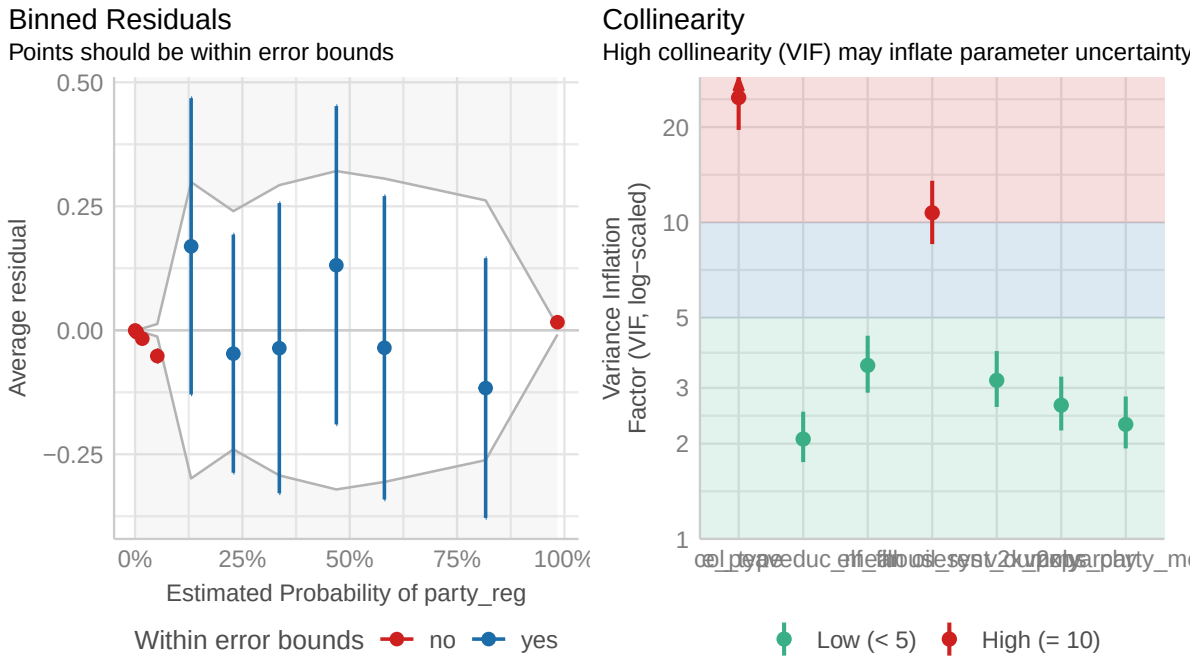

Figure 1: Model diagnostics for the preferred multinomial model.

Table 1 reveals that there are not many valid inferences to be drawn from this way of modelling the data. While democracy scores are correlated with whether the dominant regime will be autocratic or democratic dominant, this is really not a significant contribution and is indeed to be expected. Other variables that reach conventional significance levels do so when I compare both dominant regime types with non-dominant regimes. This suggests that the model is too underpowered to capture differences between autocratic and democratic dominant regimes.

With this in mind, I try and confirm this with a continuous measure of party dominance: ENPP.

```
m2 <- lm(enpp ~ v2xps_party_mean + elf_fill + col_type
         + housesys + v2x_polyarchy + oil_rent_dummy + e_peaveduc_mean,
         data = data)

m2a <- lm(enpp ~ v2xps_party_next + elf_fill + col_type
         + housesys + v2x_polyarchy + oil_rent_dummy + e_peaveduc_mean,
         data = data)

m2b <- lm(enpp ~ v2xps_party_mean + elf_fill + col_type_collapsed
         + housesys + v2x_polyarchy + oil_rent_dummy + e_peaveduc_mean,
         data = data)

m2c <- lm(enpp ~ v2xps_party_mean + elf_fill + col_type
```

```
        + housesys + v2x_polyarchy + oil_rent_perc_gdp + e_peaveduc_mean,
        data = data)


modelsummary(list("Preferred" = m2, "PSI next year imputed" = m2a,
                  "Oil rent in % of GDP" = m2b, "Colonial Legacy Collapsed" = m2c),
             stars = T, coef_rename = T, vcov = "robust",
             coef_map = c("(Intercept)", "PSI (Mean)",
                          "PSI (Next Year Imputed)", "ELF (Imputed)",
                          "Colonial Legacy [France]",
                          "Colonial Legacy [Mixed]",
                          "Colonial Legacy [None]",
                          "Colonial Legacy [Other]",
                          "Colonial Legacy (Collapsed) [France]",
                          "Colonial Legacy (Collapsed) [Other]",
                          "Electoral Rules [Plural]", "Democracy",
                          "Oil Dummy [Not equal to zero]",
                          "Oil Rents (%)",
                          "Education (Mean Imputed)"),
             output = "kableExtra") |>
  kable_styling(latex_options = c("scale_down")) |>
  column_spec(column = 2:5, width = "2.5cm")
```

Table 2: Linear model with ENPP as the dependent, all specifications

	Preferred	PSI next year imputed	Oil rent in % of GDP	Colonial Legacy Collapsed
(Intercept)	3.261*	3.242*	3.995**	3.286*
	(1.243)	(1.169)	(1.276)	(1.279)
PSI (Mean)	−3.526+		−3.988*	−3.594+
	(1.927)		(1.830)	(2.024)
PSI (Next Year Imputed)		−4.344*		
		(1.753)		
ELF (Imputed)	−0.393	−0.120	−0.415	−0.692
	(0.893)	(0.808)	(0.852)	(0.921)
Colonial Legacy [France]	0.121	−0.028		0.053
	(0.595)	(0.570)		(0.657)
Colonial Legacy [Mixed]	−1.805*	−1.760*		−1.651+
	(0.861)	(0.805)		(0.883)
Colonial Legacy [None]	−0.201	−0.510		−0.066
	(1.372)	(1.145)		(1.498)
Colonial Legacy [Other]	−0.130	−0.113		−0.107
	(0.645)	(0.644)		(0.635)
Colonial Legacy (Collapsed) [France]			−0.003	
			(0.607)	
Colonial Legacy (Collapsed) [Other]			−0.635	
			(0.547)	
Electoral Rules [Plural]	0.807	0.757	0.514	0.920
	(0.596)	(0.559)	(0.520)	(0.556)
Democracy	4.227**	4.502**	3.626*	4.140*
	(1.483)	(1.352)	(1.354)	(1.782)
Oil Dummy [Not equal to zero]	−0.489	−0.459	−0.446	
	(0.460)	(0.458)	(0.482)	
Oil Rents (%)				−0.007
				(0.029)
Education (Mean Imputed)	−0.166	−0.129	−0.152	−0.159
	(0.115)	(0.104)	(0.114)	(0.117)
Num.Obs.	37	37	37	37
R2	0.494	0.529	0.445	0.472
R2 Adj.	0.299	0.348	0.287	0.268
AIC	123.0	120.3	122.3	124.5
BIC	142.3	139.6	138.4	143.9
Log.Lik.	−49.480	−48.145	−51.168	−50.266
RMSE	0.92	0.89	0.96	0.94
Std.Errors	HC3	HC3	HC3	HC3

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

```
check_model(m2)
```

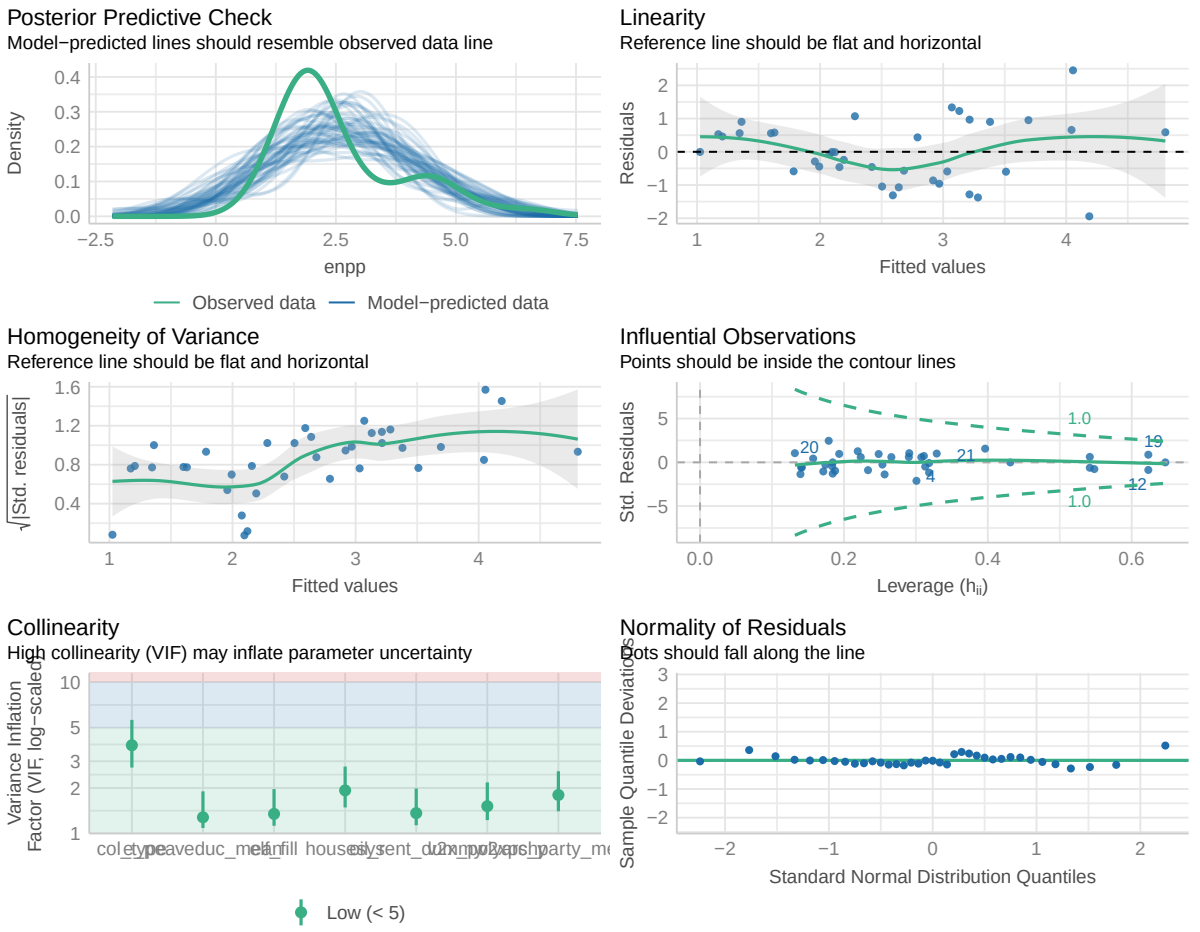

Figure 2: Model diagnostics for the preferred linear model with ENPP as the dependent variable.

When using ENPP as the dependent variable, linear models suggest similar results to those obtained using an analytical coding of the variable. While it is unsurprising, given the fact that party regime coding is based on the number of parties, it is nevertheless good to validate that approach.

Substantively, when using different operationalizations of variables, several patterns remain stable. Firstly, if anything, the preferred specification underestimates the effect of party system institutionalization on party dominance. When using different variables and especially when utilizing an alternative imputation method for missing values, the negative effect of party system institutionalization is more pronounced.

The interpretation for the latter is that the higher the party institutionalization score, the lower the expected number of parliamentary parties, that is, the higher the party dominance. This is in line with the theoretical expectations.

Since it appears that the modelling capabilities are restricted, I also evaluate a simpler logistic model, where the coding does not differentiate between autocratic- and democratic- dominant regimes, that is, the countries either are (1) dominant or (0) non-dominant regimes.

```
m3 <- glm(party_reg_dummy ~ v2xps_party_mean + elf_fill + col_type
          + housesys + v2x_polyarchy + oil_rent_dummy + e_peaveduc_mean,
          data = data, family = "binomial")

m3a <- glm(party_reg_dummy ~ v2xps_party_next + elf_fill + col_type
          + housesys + v2x_polyarchy + oil_rent_dummy + e_peaveduc_mean,
          data = data, family = "binomial")

m3b <- glm(party_reg_dummy ~ v2xps_party_mean + elf_fill + col_type_collapsed
          + housesys + v2x_polyarchy + oil_rent_dummy + e_peaveduc_mean,
          data = data, family = "binomial")

m3c <- glm(party_reg_dummy ~ v2xps_party_mean + elf_fill + col_type
          + housesys + v2x_polyarchy + oil_rent_perc_gdp + e_peaveduc_mean,
          data = data, family = "binomial")

modelsummary(list("Preferred" = m3, "PSI next year imputed" = m3a,
                  "Oil rent in % of GDP" = m3b, "Colonial Legacy Collapsed" = m3c),
             stars = T, coef_rename = T, vcov = "robust",
             coef_map = c("(Intercept)", "PSI (Mean)",
                          "PSI (Next Year Imputed)", "ELF (Imputed)",
                          "Colonial Legacy [France]",
                          "Colonial Legacy [Mixed]",
                          "Colonial Legacy [None]",
                          "Colonial Legacy [Other]",
                          "Colonial Legacy (Collapsed) [France]",
                          "Colonial Legacy (Collapsed) [Other]",
                          "Electoral Rules [Plural]", "Democracy",
                          "Oil Dummy [Not equal to zero]",
                          "Oil Rents (%)",
                          "Education (Mean Imputed)"),
             output = "kableExtra") |>
  kable_styling(latex_options = c("scale_down")) |>
  column_spec(column = 2:5, width = "2.5cm")
```

The results from a simplified categorization of party domination mostly reinforce previous results, but also produces unorthodox estimates for the effect of "mixed" colonial legacies. When the coefficients are exponentiated, those countries that have had such legacies, are 2543369 times more likely to be dominant than non-dominant. This is, no doubt a consequence of small sample size and a subsequent "small cells" problem. When I cross-reference the two variables, the issue is obvious: all three countries with a "mixed" colonial legacy are also autocratic dominant.

Table 3: Logistic model with party dominance dummy as the dependent, all specifications

	Preferred	PSI next year imputed	Oil rent in % of GDP	Colonial Legacy Collapsed
(Intercept)	−4.482	−3.159	−3.567	−5.367
	(6.785)	(5.136)	(5.070)	(6.482)
PSI (Mean)	−10.940		−10.330	−8.804
	(12.150)		(12.265)	(10.238)
PSI (Next Year Imputed)		−27.059		
		(19.985)		
ELF (Imputed)	−3.395	−2.868	−3.439	−2.098
	(5.220)	(5.693)	(4.698)	(3.911)
Colonial Legacy [France]	−0.033	−2.592		0.526
	(2.740)	(2.625)		(2.445)
Colonial Legacy [Mixed]	14.749***	14.234***		14.568***
	(3.507)	(4.229)		(3.005)
Colonial Legacy [None]	−1.679	−6.566		−1.477
	(4.228)	(4.462)		(3.709)
Colonial Legacy [Other]	3.948	3.238		4.031
	(4.729)	(4.381)		(4.131)
Colonial Legacy (Collapsed) [France]			−0.179	
			(2.528)	
Colonial Legacy (Collapsed) [Other]			2.601	
			(2.922)	
Electoral Rules [Plural]	5.073	6.674	4.201+	4.767
	(3.438)	(5.182)	(2.436)	(3.067)
Democracy	18.087*	28.244+	17.324*	16.778*
	(7.801)	(16.388)	(7.694)	(6.857)
Oil Dummy [Not equal to zero]	0.759	2.473	0.726	
	(2.513)	(4.021)	(2.270)	
Oil Rents (%)				−0.170
				(0.260)
Education (Mean Imputed)	−0.141	0.086	−0.145	−0.166
	(0.491)	(0.376)	(0.462)	(0.451)
Num.Obs.	38	38	38	38
AIC	47.6	41.6	44.8	47.2
BIC	65.6	59.6	59.5	65.3
Log.Lik.	−12.787	−9.813	−13.386	−12.622
F	8.434	3.071	1.160	10.218
RMSE	0.34	0.28	0.34	0.34
Std.Errors	HC3	HC3	HC3	HC3

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

```
kable(table(data$party_reg_dummy, data$col_type))
```

Table 4: Cross-tabulation of party dominance and colonial legacies variables

	Britain	France	Mixed	None	Other
Dominant	9	7	0	1	4
Not Dominant	7	7	3	1	3

This is also a consequence of my decision to model all possible categories of colonial legacies. However, there is no reason to think this affects the coefficients on other variables, as the model with a "collapsed" version of the colonial legacies variable behaves fine.

```
check_model(m3)
```

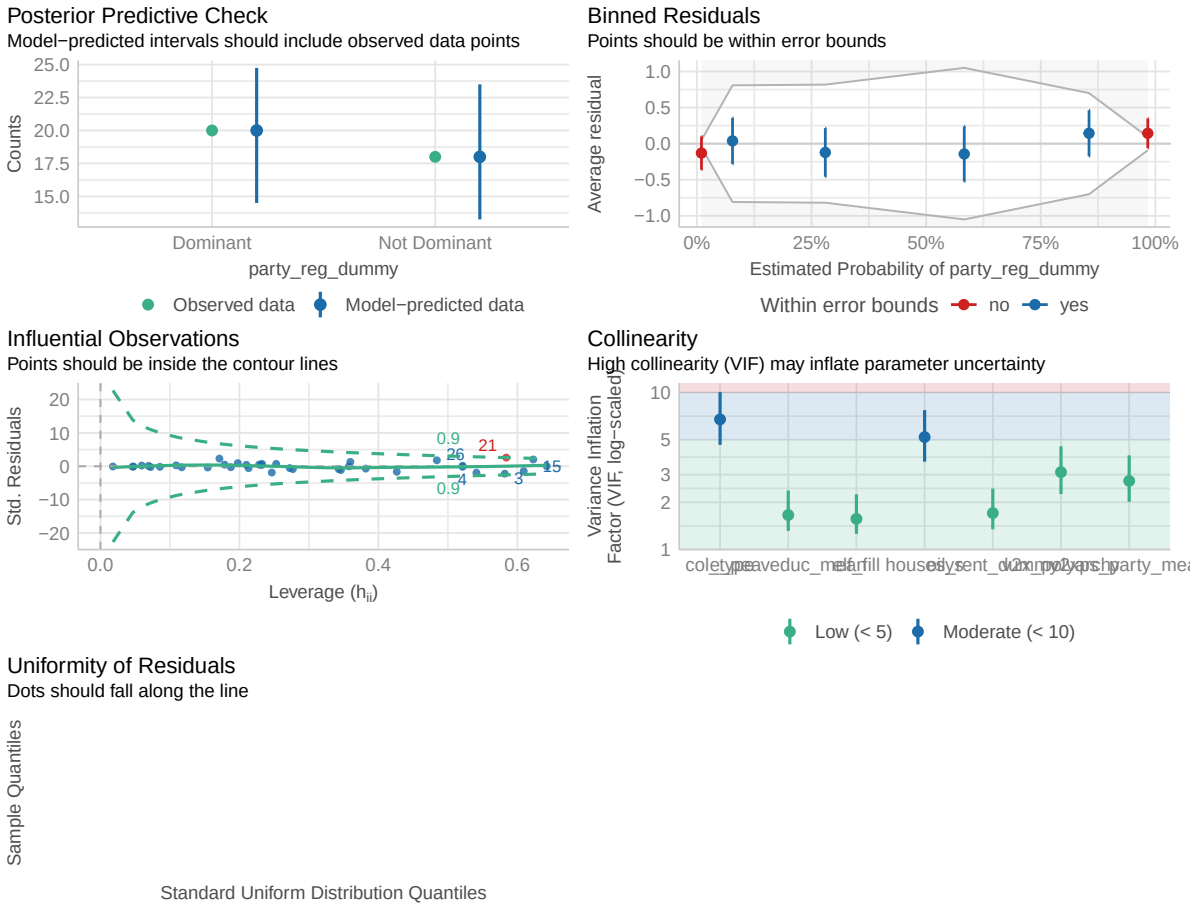

Figure 3: Model diagnostics for the preferred logistic model with party dominance dummy as the dependent variable.