intermediate analysis

Sean McSweeney

IRI Voting Integrity Analysis

Exploratory Analysis

Missing data per variable

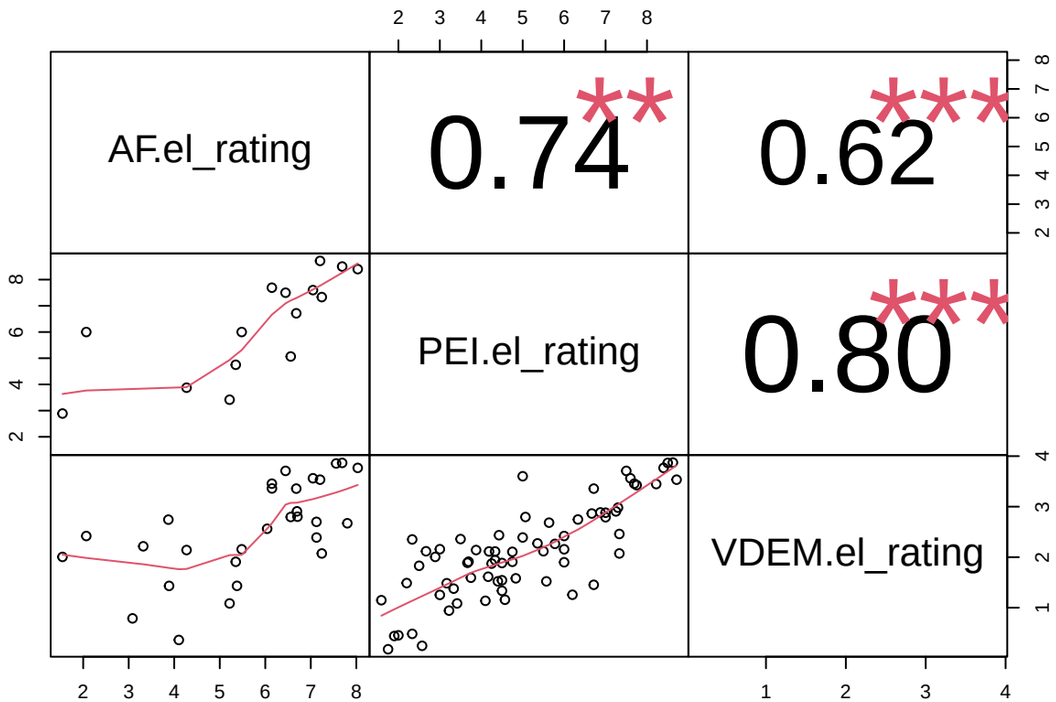
```
perc_complete <- bind_rows(map(merged_datasets3, ~mean(is.na(.))))
perc_complete <- perc_complete %>%
  pivot_longer(cols = everything()) %>%
  mutate(`%complete` = round(1 - value, digits = 2)) %>%
  select(!value) %>%
  arrange(-`%complete`)

perc_complete %>% write_csv("CleanedMergedData/perc_complete.csv")

stargazer(perc_complete, summary = FALSE, rownames = FALSE, float = FALSE)
```

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu % Date and time: Tue, May 17, 2022 - 16:57:59

name	%complete
country_year	1
country	1
year	1
v2xel_elecpres	0.92
v2xel_elecparl	0.92
VDEM.el_rating	0.92
VDEM.EMB_cap	0.92
VDEM.elec_viol	0.92
VDEM.access_to_public_services_social_group	0.92
VDEM.access_to_public_services_pol_group	0.92
VDEM.is_an_ethnic_group_the_most_powerful_regime_support_group	0.92
VDEM.is_an_ethnic_group_the_most_powerful_regime_duration_group	0.92
VDEM.international_monitors_present	0.91
nelda11	0.88
nelda3	0.88
nelda5	0.88
nelda45	0.87
nelda4	0.87
VDEM.turnout	0.86
nelda18	0.86
WB.FDI	0.86
WB.urban_pop	0.86
VDEM.regime_support_by_an_ethnic_group	0.86
VDEM.VAP_turnout	0.85
nelda17	0.85
nelda29	0.85
FH.fh_total	0.85
WB.GDP_ppc	0.85
literacy_imputed	0.84
WB.inflation	0.81
VDEM.yrs_education	0.78
nelda46	0.77
POLITY5.fragment	0.76
POLITY5.polity2	0.76
POLITY5.durable	0.76
VDEM.el_rating.lagged	0.76
POLITY5.parcomp	0.74
press_freedom	0.7
electexec	0.52
electboth	0.52
electleg	0.51
PEI.EMB_cap	0.25
PEI.protest	0.25

correlation analysis for different integrity measures:

```
chart.Correlation(select(merged_datasets3, AF.el_rating, PEI.el_rating, VDEM.el_rating),
                  histogram = FALSE)
```

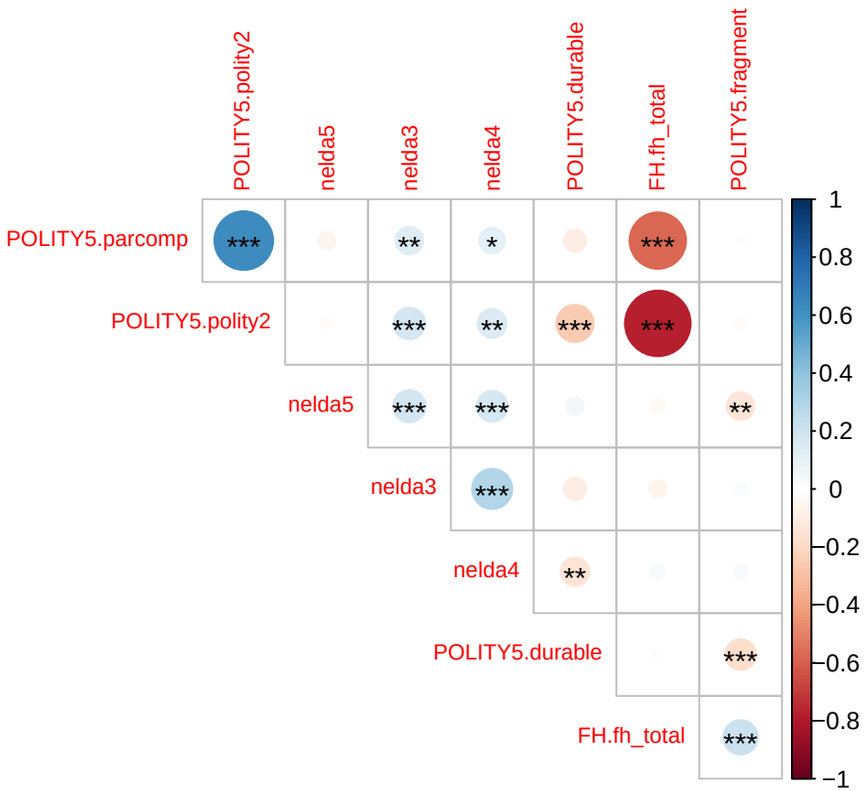


Correlations between causes independent variables by category:

Governance

```
gov_indc <- c("nelda3", "nelda4", "nelda5", "POLITY5.parcomp", "FH.fh_total", "POLITY5.fragment",
              "POLITY5.polity2", "POLITY5.durable")
c_gov <- rcorr(as.matrix(select(merged_datasets3, unlist(gov_indc))))
```

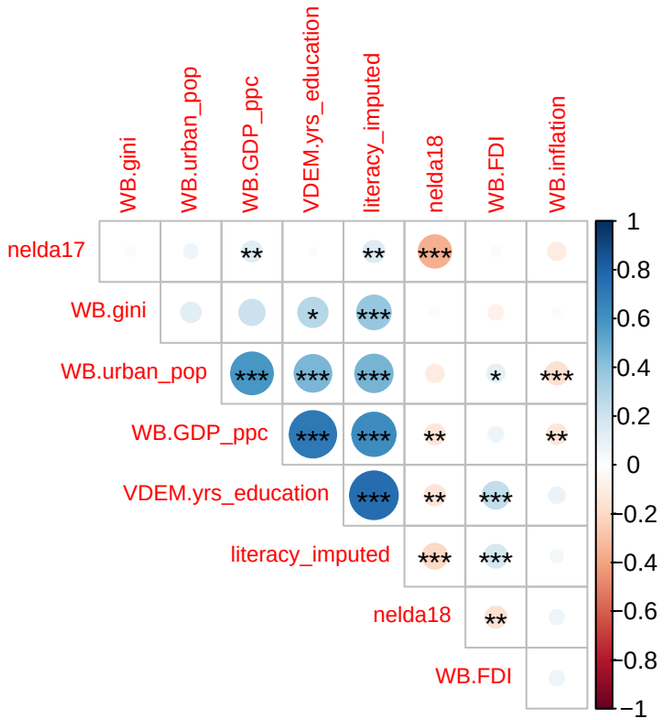
```
corrplot(c_gov$r,  method = "circle", number.cex = 0.75, type = "upper", tl.cex = 0.75,
         order = "hclust", p.mat = c_gov$P, sig.level = c(.01, .05, 0.1),  diag = FALSE,
         insig = "label_sig", pch.cex = 1)
```



Socioeconomic

```
socioecon_indc <- c("nelda17", "nelda18", "WB.GDP_ppc", "WB.FDI", "VDEM.yrs_education",
                    "WB.inflation", "WB.gini", "WB.urban_pop", "literacy_imputed")
c_socio <- rcorr(as.matrix(select(merged_datasets3, unlist(socioecon_indc))))
```

```
corrplot(c_socio$r,  method = "circle", number.cex = 0.75, type = "upper", tl.cex = 0.75,
         order = "hclust", p.mat = c_socio$P, sig.level = c(.01, .05, 0.1), diag= FALSE,
         insig = "label_sig", pch.cex = 1, mar=c(0,0,2,0))
```

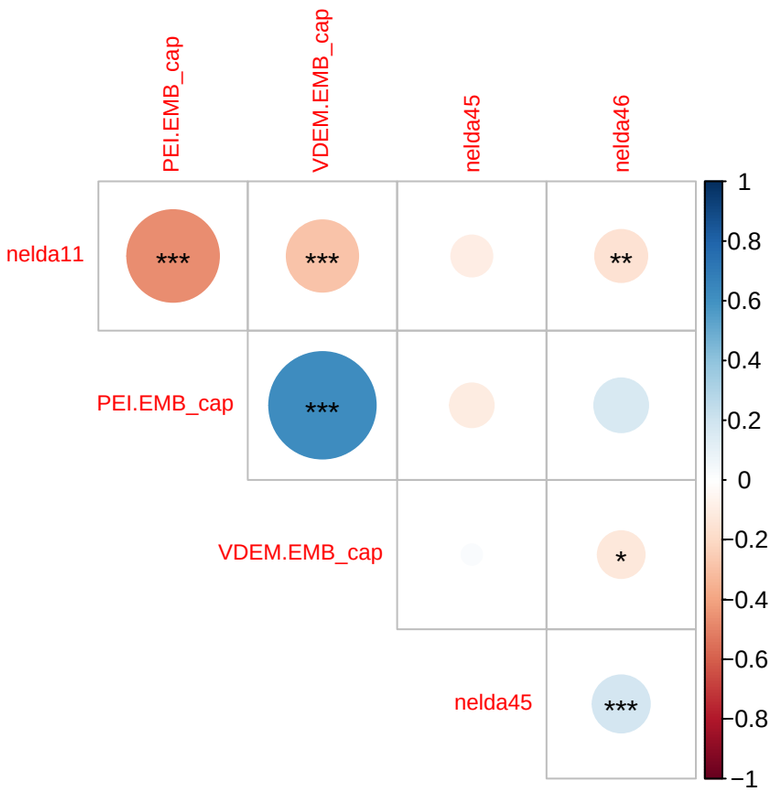


Media

press_freedom

EMB

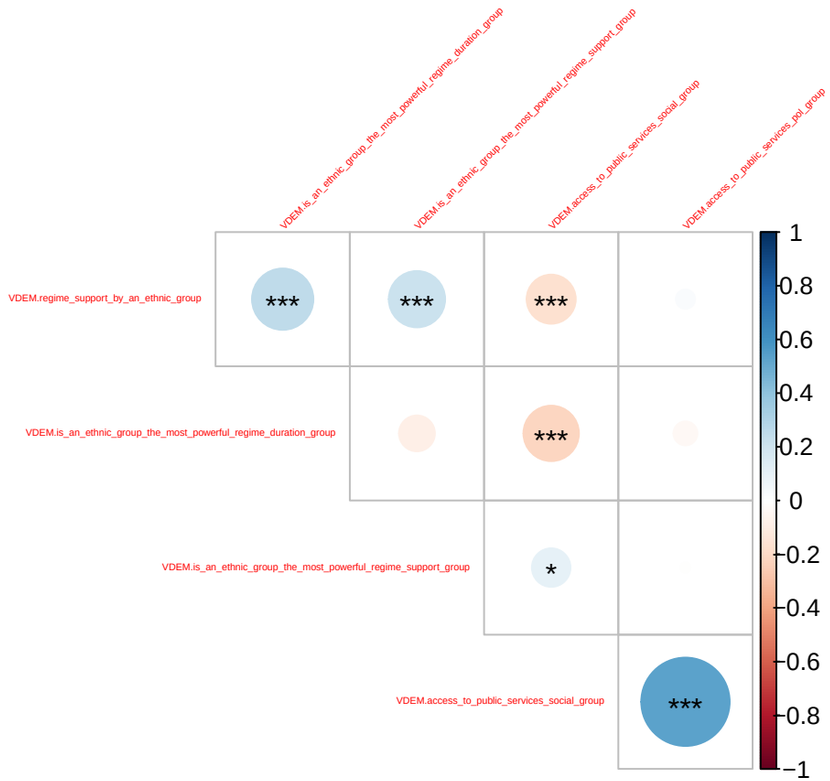
```r
emb_indc <- c("PEI.EMB_cap","VDEM.EMB_cap", "nelda11", "nelda45", "nelda46")
c_emb <- rcorr(as.matrix(select(merged_datasets3, unlist(emb_indc))))
corrplot(c_emb$r,  method = "circle", number.cex = 0.75, type = "upper", tl.cex = 0.75,
         order = "hclust", p.mat = c_emb$P, sig.level = c(.01, .05, 0.1), diag= FALSE,
         insig = "label_sig", pch.cex = 1)
```



pre-election violence

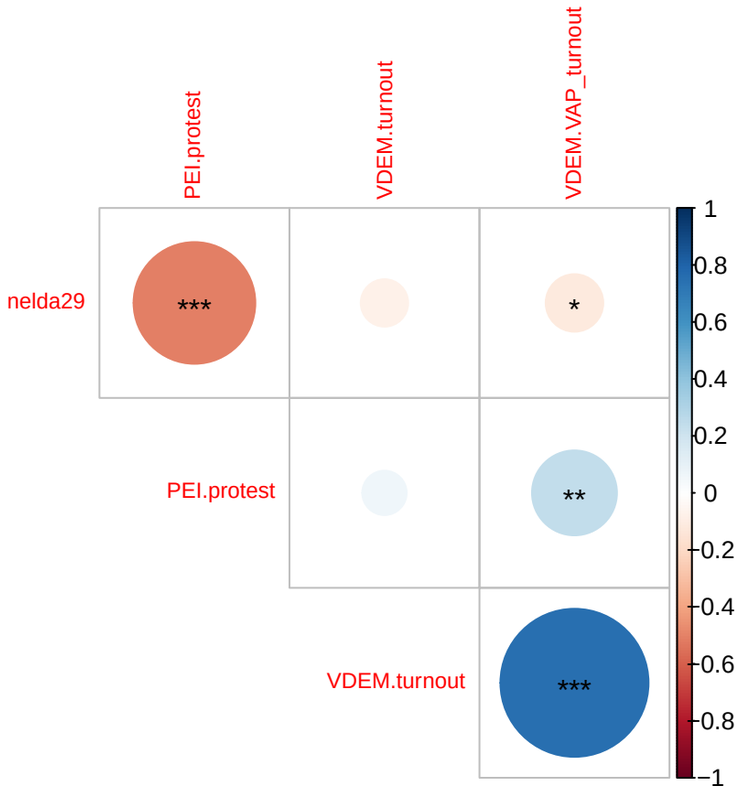
ethnic divisions

```
ethnic_indc <- c("VDEM.access_to_public_services_social_group",
                 "VDEM.access_to_public_services_pol_group",
                 "VDEM.regime_support_by_an_ethnic_group",
                 "VDEM.is_an_ethnic_group_the_most_powerful_regime_support_group",
                 "VDEM.is_an_ethnic_group_the_most_powerful_regime_duration_group")
c_ethnic <- rcorr(as.matrix(select(merged_datasets3, unlist(ethnic_indc))))
corrplot(c_ethnic$r,  method = "circle", number.cex = 0.75, type = "upper", tl.cex = 0.3,
         order = "hclust", p.mat = c_ethnic$P, sig.level = c(.01, .05, 0.1), diag= FALSE,
         insig = "label_sig", pch.cex = 1, tl.srt = 45)
```



Consequences

```
conseq_indc <- c("PEI.protest", "nelda29", "VDEM.turnout", "VDEM.VAP_turnout")
c_conseq <- rcorr(as.matrix(select(merged_datasets3, unlist(conseq_indc))))
corrplot(c_conseq$r,  method = "circle", number.cex = 0.75, type = "upper", tl.cex = 0.75,
         order = "hclust", p.mat = c_conseq$P, sig.level = c(.01, .05, 0.1), diag= FALSE,
         insig = "label_sig", pch.cex = 1)
```



individually regressing all variables on VDEM.el_rating

Governance

```
lm.nelda3 <- lm(VDEM.el_rating ~ nelda3, merged_datasets3)
lm.nelda4 <- lm(VDEM.el_rating ~ nelda4, merged_datasets3)
lm.nelda5 <- lm(VDEM.el_rating ~ nelda5, merged_datasets3)
lm.POLITY5.parcomp <- lm(VDEM.el_rating ~ POLITY5.parcomp, merged_datasets3)
lm.FH.fh_total <- lm(VDEM.el_rating ~ FH.fh_total, merged_datasets3)
```

```
lm.POLITY5.fragment <- lm(VDEM.el_rating ~ POLITY5.fragment, merged_datasets3)
lm.POLITY5.polity2 <- lm(VDEM.el_rating ~ POLITY5.polity2, merged_datasets3)
lm.POLITY5.durable <- lm(VDEM.el_rating ~ POLITY5.durable, merged_datasets3)

stargazer(title = "Governance indicators", lm.nelda3, lm.nelda4, lm.nelda5, lm.POLITY5.parcomp,
          lm.FH.fh_total, lm.POLITY5.fragment, lm.POLITY5.polity2,
          lm.POLITY5.durable, font.size= "small", single.row = TRUE, column.sep.width = "1pt", no.space = TRUE)
```

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu % Date and time: Tue, May 17, 2022 - 16:58:04

Table 1: Governance indicators

	Dependent variable:							
	VDEM.el_rating							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
nelda3	0.796** (0.379)							
nelda4		0.517 (0.353)						
nelda5			0.355 (0.466)					
POLITY5.parcomp				0.427*** (0.048)				
FH.fh_total					−0.246*** (0.013)			
POLITY5.fragment						0.197 (0.158)		
POLITY5.polity2							0.129*** (0.009)	
POLITY5.durable								−0.003 (0.005)
Constant	1.371*** (0.375)	1.651*** (0.349)	1.800*** (0.463)	0.821*** (0.154)	4.287*** (0.116)	2.118*** (0.064)	1.893*** (0.048)	2.162*** (0.091)
Observations	257	256	257	211	244	217	217	217
R^2	0.017	0.008	0.002	0.279	0.607	0.007	0.495	0.001
Adjusted R^2	0.013	0.004	−0.002	0.276	0.605	0.003	0.492	−0.004
Residual Std. Error	0.918 (df = 255)	0.922 (df = 254)	0.925 (df = 255)	0.798 (df = 209)	0.575 (df = 242)	0.933 (df = 215)	0.666 (df = 215)	0.936 (df = 215)
F Statistic	4.408** (df = 1; 255)	2.137 (df = 1; 254)	0.580 (df = 1; 255)	80.895*** (df = 1; 209)	373.699*** (df = 1; 242)	1.556 (df = 1; 215)	210.558*** (df = 1; 215)	0.219 (df = 1; 215)
Note:								*p<0.1; **p<0.05; ***p<0.01

Socioeconomic

```
lm.nelda17 <- lm(VDEM.el_rating ~ nelda17, merged_datasets3)
lm.nelda18 <- lm(VDEM.el_rating ~ nelda18, merged_datasets3)
lm.WB.GDP_ppc <- lm(VDEM.el_rating ~ WB.GDP_ppc, merged_datasets3)
```

```
lm.WB.FDI <- lm(VDEM.el_rating ~ WB.FDI, merged_datasets3)
lm.VDEM.yrs_education <- lm(VDEM.el_rating ~ VDEM.yrs_education, merged_datasets3)
lm.WB.inflation <- lm(VDEM.el_rating ~ WB.inflation, merged_datasets3)
lm.WB.gini <- lm(VDEM.el_rating ~ WB.gini, merged_datasets3)
lm.WB.urban_pop <- lm(VDEM.el_rating ~ WB.urban_pop, merged_datasets3)
lm.literacy_imputed <- lm(VDEM.el_rating ~ literacy_imputed, merged_datasets3)

stargazer(title = "Socioeconomic indicators", lm.nelda17, lm.nelda18, lm.WB.GDP_ppc, lm.WB.FDI,
          lm.VDEM.yrs_education, lm.WB.inflation, lm.WB.gini,
          lm.WB.urban_pop, lm.literacy_imputed, font.size= "small", single.row = TRUE, column.sep.width = "1pt", no.space = TRUE)
```

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu % Date and time: Tue, May 17, 2022 - 16:58:04

Table 2: Socioeconomic indicators

	Dependent variable:								
	VDEM.el_rating								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
nelda17	−0.136 (0.121)								
nelda18		−0.143 (0.138)							
WB.GDP_ppc			0.00000 (0.00002)						
WB.FDI				−0.000 (0.000)					
VDEM.yrs_education					0.005 (0.026)				
WB.inflation						−0.001 (0.011)			
WB.gini							0.053*** (0.015)		
WB.urban_pop								0.001 (0.003)	
literacy_imputed									−0.0003 (0.003)
Constant	2.215*** (0.074)	2.190*** (0.066)	2.192*** (0.076)	2.202*** (0.066)	2.126*** (0.142)	2.232*** (0.087)	−0.044 (0.677)	2.117*** (0.156)	2.224*** (0.192)
Observations	250	252	244	248	242	235	42	248	241
R^2	0.005	0.004	0.0001	0.002	0.0002	0.00002	0.229	0.001	0.00004
Adjusted R^2	0.001	0.0003	−0.004	−0.002	−0.004	−0.004	0.209	−0.003	−0.004
Residual Std. Error	0.925 (df = 248)	0.924 (df = 250)	0.930 (df = 242)	0.932 (df = 246)	0.893 (df = 240)	0.912 (df = 233)	0.833 (df = 40)	0.932 (df = 246)	0.932 (df = 239)
F Statistic	1.256 (df = 1; 248)	1.076 (df = 1; 250)	0.014 (df = 1; 242)	0.594 (df = 1; 246)	0.039 (df = 1; 240)	0.006 (df = 1; 233)	11.856*** (df = 1; 40)	0.190 (df = 1; 246)	0.009 (df = 1; 239)
Note:									*p<0.1; **p<0.05; ***p<0.01

Media

```
lm.press_freedom <- lm(VDEM.el_rating ~ press_freedom, merged_datasets3)

stargazer(title = "Press indicators", lm.press_freedom)
```

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu % Date and time: Tue, May 17, 2022 - 16:58:05

Table 3: Press indicators

	Dependent variable:
	VDEM.el_rating
press_freedom	-0.038^{***}
	(0.003)
Constant	3.352^{***}
	(0.119)
Observations	209
R^2	0.368
Adjusted R^2	0.365
Residual Std. Error	0.724 (df = 207)
F Statistic	120.732^{***} (df = 1; 207)
Note:	*p<0.1; **p<0.05; ***p<0.01

EMB

```
lm.VDEM.EMB_cap <- lm(VDEM.el_rating ~ VDEM.EMB_cap, merged_datasets3)
lm.nelda11 <- lm(VDEM.el_rating ~ nelda11, merged_datasets3)
lm.nelda45 <- lm(VDEM.el_rating ~ nelda45, merged_datasets3)
lm.nelda46 <- lm(VDEM.el_rating ~ nelda46, merged_datasets3)

stargazer(title = "EMB indicators", lm.VDEM.EMB_cap, lm.nelda11, lm.nelda45, lm.nelda46, single.row = TRUE)
```

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu % Date and time: Tue, May 17, 2022 - 16:58:05

Table 4: EMB indicators

	Dependent variable:			
	VDEM.el_rating			
	(1)	(2)	(3)	(4)
VDEM.EMB_cap	0.600*** (0.055)			
nelda11		−0.761*** (0.108)		
nelda45			0.280 (0.188)	
nelda46				0.116 (0.132)
Constant	0.913*** (0.119)	2.609*** (0.084)	1.898*** (0.178)	2.091*** (0.110)
Observations	283	257	255	229
R^2	0.300	0.163	0.009	0.003
Adjusted R^2	0.298	0.160	0.005	−0.001
Residual Std. Error	0.774 (df = 281)	0.847 (df = 255)	0.923 (df = 253)	0.924 (df = 227)
F Statistic	120.454*** (df = 1; 281)	49.681*** (df = 1; 255)	2.214 (df = 1; 253)	0.763 (df = 1; 227)
Note:				*p<0.1; **p<0.05; ***p<0.01

ethnic divisions

```
lm.VDEM.access_to_public_services_social_group <- lm(VDEM.el_rating ~ VDEM.access_to_public_services_social_group, merged_datasets3)
lm.VDEM.access_to_public_services_pol_group <- lm(VDEM.el_rating ~ VDEM.access_to_public_services_pol_group, merged_datasets3)
lm.VDEM.regime_support_by_an_ethnic_group <- lm(VDEM.el_rating ~ VDEM.regime_support_by_an_ethnic_group, merged_datasets3)
lm.VDEM.is_an_ethnic_group_the_most_powerful_regime_support_group <- lm(VDEM.el_rating ~ VDEM.is_an_ethnic_group_the_most_powerful_regime_support_group, merged_datasets3)
lm.VDEM.is_an_ethnic_group_the_most_powerful_regime_duration_group <- lm(VDEM.el_rating ~ VDEM.is_an_ethnic_group_the_most_powerful_regime_duration_group, merged_datasets3)

stargazer(title = "Ethnic indicators", lm.VDEM.access_to_public_services_social_group, lm.VDEM.access_to_public_services_pol_group, lm.VDEM.regime_support_by_an_ethnic_group,
          lm.VDEM.is_an_ethnic_group_the_most_powerful_regime_support_group, lm.VDEM.is_an_ethnic_group_the_most_powerful_regime_duration_group,
          single.row = TRUE)
```

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu % Date and time: Tue, May 17, 2022 - 16:58:05

Consequences

Table 5: Ethnic indicators

	Dependent variable:				
	VDEM.el_rating				
	(1)	(2)	(3)	(4)	(5)
VDEM.access_to_public_services_social_group	0.299*** (0.054)				
VDEM.access_to_public_services_pol_group		0.588*** (0.052)			
VDEM.regime_support_by_an_ethnic_group			0.337 (0.268)		
VDEM.is_an_ethnic_group_the_most_powerful_regime_support_group				−0.002 (0.225)	
VDEM.is_an_ethnic_group_the_most_powerful_regime_duration_group					−0.308* (0.180)
Constant	2.034*** (0.054)	2.005*** (0.047)	2.031*** (0.090)	2.119*** (0.057)	2.150*** (0.058)
Observations	283	283	266	283	283
R^2	0.098	0.316	0.006	0.00000	0.010
Adjusted R^2	0.095	0.314	0.002	−0.004	0.007
Residual Std. Error	0.879 (df = 281)	0.765 (df = 281)	0.919 (df = 264)	0.926 (df = 281)	0.921 (df = 281)
F Statistic	30.492*** (df = 1; 281)	130.009*** (df = 1; 281)	1.584 (df = 1; 264)	0.0001 (df = 1; 281)	2.909* (df = 1; 281)
Note:					*p<0.1; **p<0.05; ***p<0.01

```
lm.PEI.protest <- lm(VDEM.el_rating ~ PEI.protest, merged_datasets3)
lm.nelda29 <- lm(VDEM.el_rating ~ nelda29, merged_datasets3)
lm.VDEM.turnout <- lm(VDEM.el_rating ~ VDEM.turnout, merged_datasets3)
lm.VDEM.VAP_turnout <- lm(VDEM.el_rating ~ VDEM.VAP_turnout, merged_datasets3)

stargazer(title = "Consequnces indicators", lm.PEI.protest, lm.VDEM.turnout, lm.VDEM.VAP_turnout)
```

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu % Date and time: Tue, May 17, 2022 - 16:58:05

Table 6: Consequnces indicators

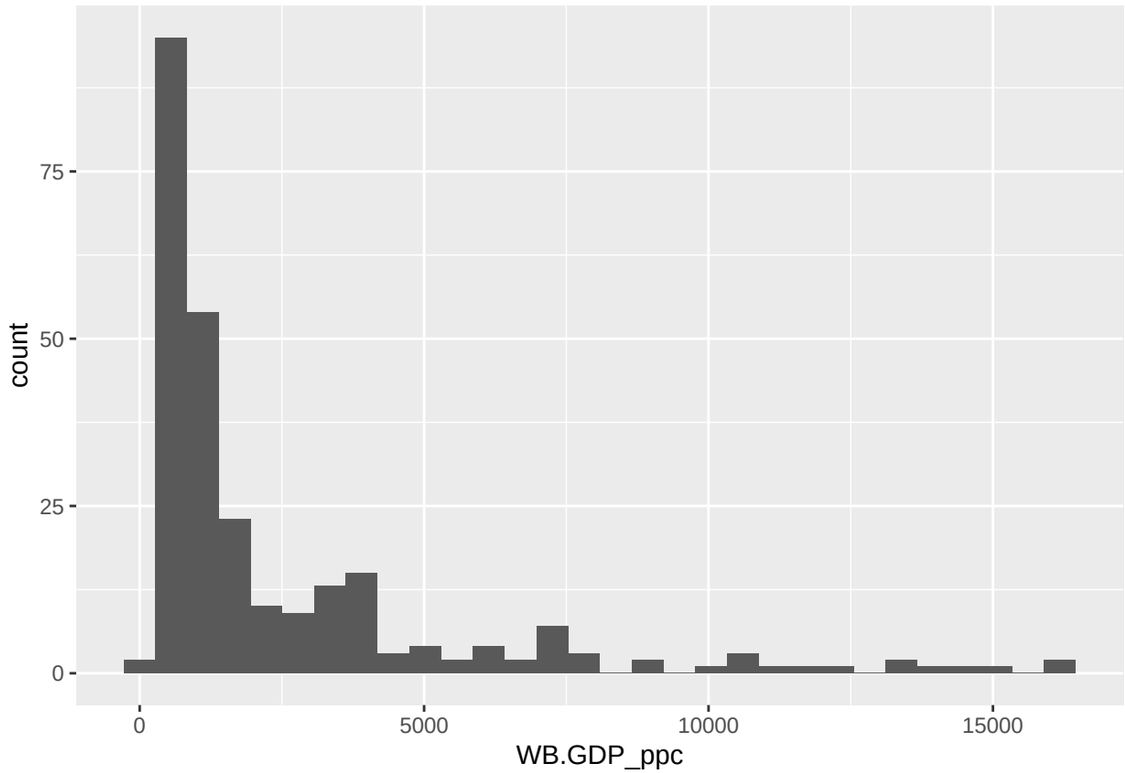
	Dependent variable:		
	VDEM.el_rating		
	(1)	(2)	(3)
PEI.protest	0.030***		
	(0.006)		
VDEM.turnout		0.006*	
		(0.003)	
VDEM.VAP_turnout			0.005*
			(0.003)
Constant	0.588*	1.788***	1.884***
	(0.318)	(0.223)	(0.166)
Observations	73	267	264
R^2	0.273	0.011	0.011
Adjusted R^2	0.262	0.007	0.008
Residual Std. Error	0.785 (df = 71)	0.920 (df = 265)	0.918 (df = 262)
F Statistic	26.598*** (df = 1; 71)	2.838* (df = 1; 265)	3.033* (df = 1; 262)
Note:			*p<0.1; **p<0.05; ***p<0.01

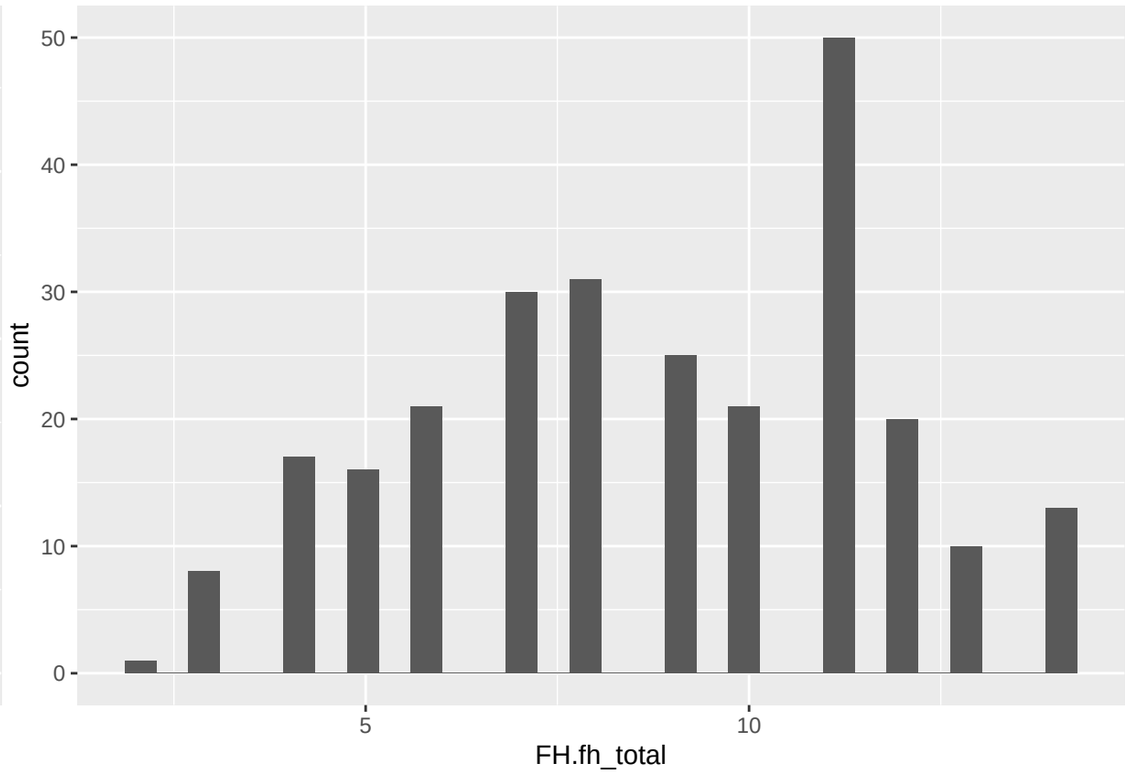
Trying models!

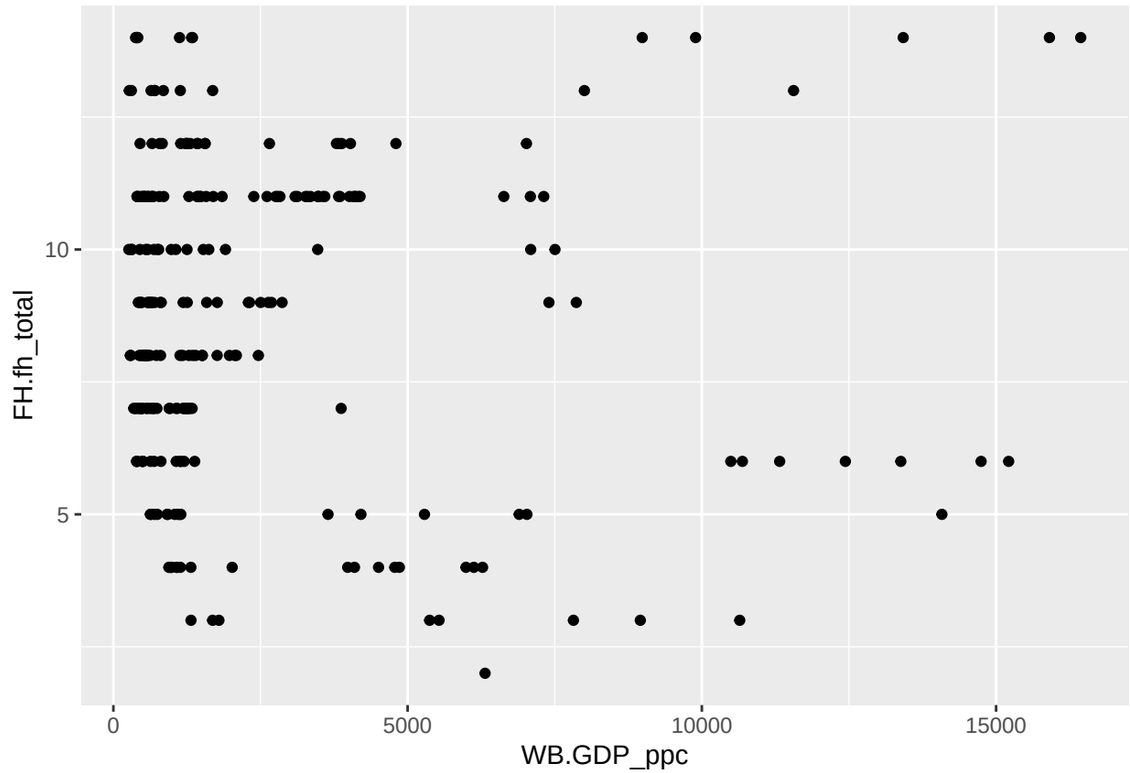
Subgrouping income/democracy

Dividing the country-year's into three different development levels by income per capita. Also dividing country-year's into different levels of democracy/autocracy.
Income per capita levels:
4: (3140, 16400]
3: (1220, 3140]
2: (627, 1220]
1: [262, 627]

FreedomHouse levels:
4: (11, 14]
3: (9, 11]
2: (7, 9]
1: [2, 7]

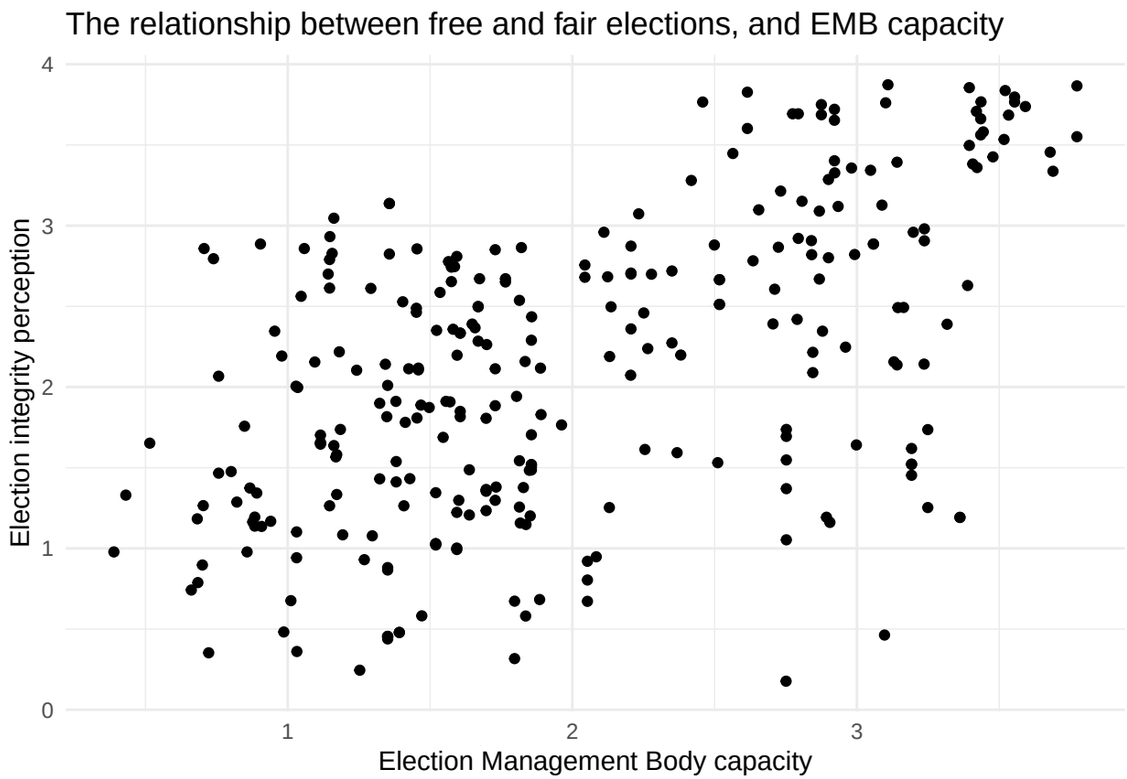

count
75
50
25
0
0
5000
10000
15000
WB.GDP_ppc


count
50
40
30
20
10
0
5
10
FH.fh_total



EMB capacity models

```
ggplot(merged_datasets3) +
  geom_point(aes(VDEM.EMB_cap, VDEM.el_rating)) +
  theme_minimal() +
  labs(x = "Election Management Body capacity",
       y = "Election integrity perception",
       title = "The relationship between free and fair elections, and EMB capacity")
```




```
stargazer(lm(VDEM.el_rating ~ VDEM.EMB_cap + income_level, merged_datasets3))
```

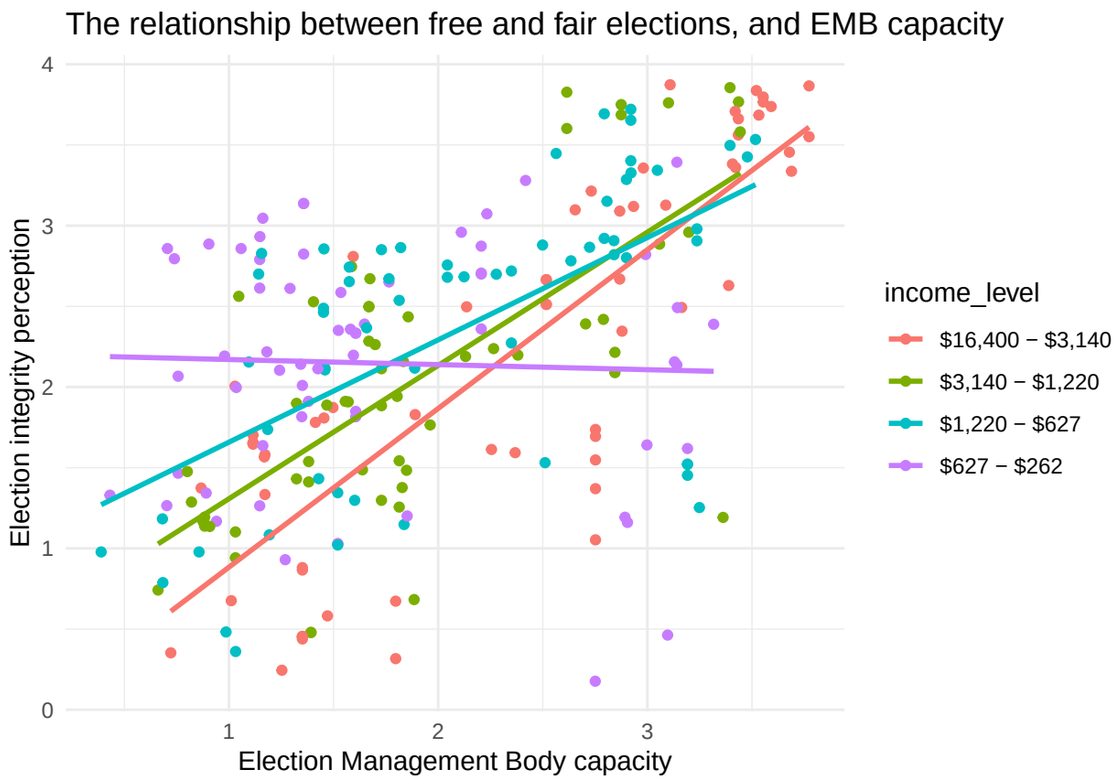
% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu % Date and time: Tue, May 17, 2022 - 16:58:08

```
ggplot(drop_na(merged_datasets3, income_level), aes(VDEM.EMB_cap, VDEM.el_rating, color = income_level)) +
  geom_point() +
  theme_minimal() +
  labs(x = "Election Management Body capacity",
       y = "Election integrity perception",
       title = "The relationship between free and fair elections, and EMB capacity") +
  geom_smooth(method = "lm", se = FALSE)
```

```
## `geom_smooth()` using formula 'y ~ x'
```

Table 7:

	Dependent variable:
	VDEM.el_rating
VDEM.EMB_cap	0.642***
	(0.058)
1,220	0.139
	(0.139)
627	0.311**
	(0.136)
262	0.354**
	(0.140)
Constant	0.697***
	(0.165)
Observations	244
R^2	0.351
Adjusted R^2	0.340
Residual Std. Error	0.754 (df = 239)
F Statistic	32.263*** (df = 4; 239)
Note:	*p<0.1; **p<0.05; ***p<0.01




```
stargazer(lm(VDEM.el_rating ~ VDEM.EMB_cap + fh_level, merged_datasets3))
```

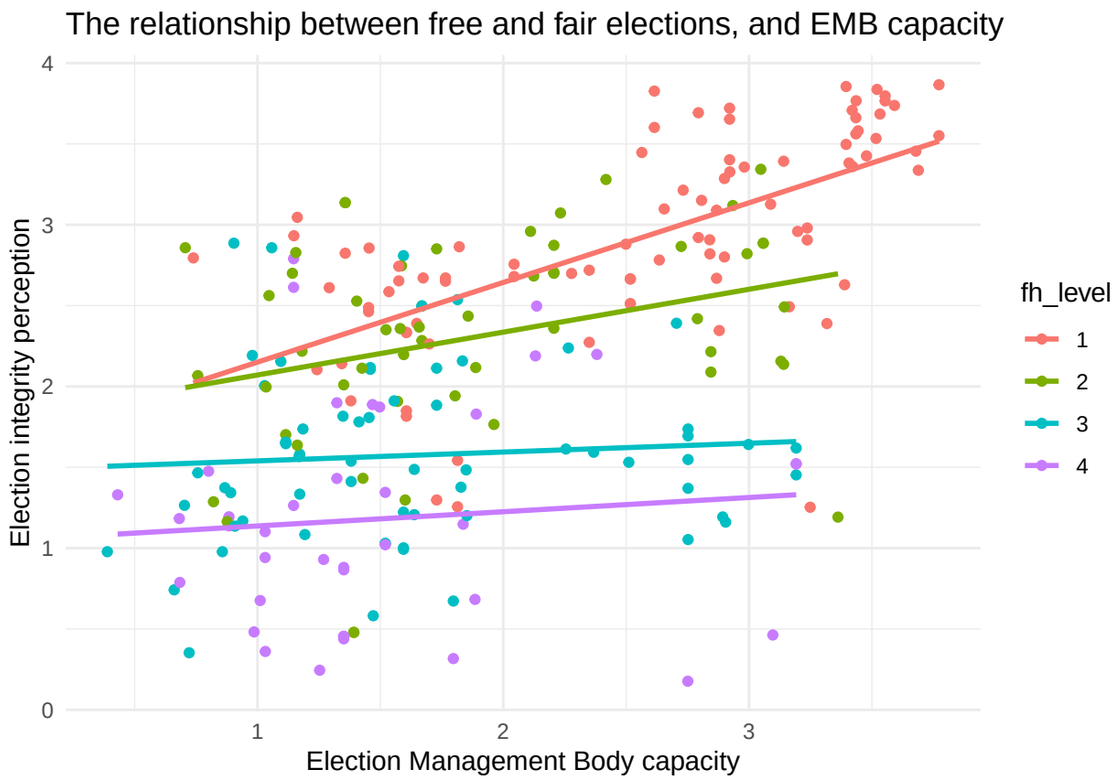
% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu % Date and time: Tue, May 17, 2022 - 16:58:09

```
ggplot(drop_na(merged_datasets3, fh_level), aes(VDEM.EMB_cap, VDEM.el_rating, color = fh_level)) +
  geom_point() +
  theme_minimal() +
  labs(x = "Election Management Body capacity",
       y = "Election integrity perception",
       title = "The relationship between free and fair elections, and EMB capacity") +
  geom_smooth(method = "lm", se = FALSE)
```

```
## `geom_smooth()` using formula 'y ~ x'
```

Table 8:

	Dependent variable:
	VDEM.el_rating
VDEM.EMB_cap	0.288^{***}
	(0.050)
fh_level2	-0.420^{***}
	(0.107)
fh_level3	-1.079^{***}
	(0.107)
fh_level4	-1.433^{***}
	(0.125)
Constant	2.182^{***}
	(0.144)
Observations	244
R^2	0.599
Adjusted R^2	0.592
Residual Std. Error	0.585 (df = 239)
F Statistic	89.148^{***} (df = 4; 239)
Note:	*p<0.1; **p<0.05; ***p<0.01



lasso modeling

fixed effects

create panel dataset

```
merged.p <- pdata.frame(select(merged_datasets3, !country_year), index = c("country", "year"))
fixedeff <- plm(VDEM.el_rating ~ VDEM.EMB_cap, data = merged.p, model = "within")
stargazer(fixedeff, lm(VDEM.el_rating ~ VDEM.EMB_cap, merged_datasets3))
```

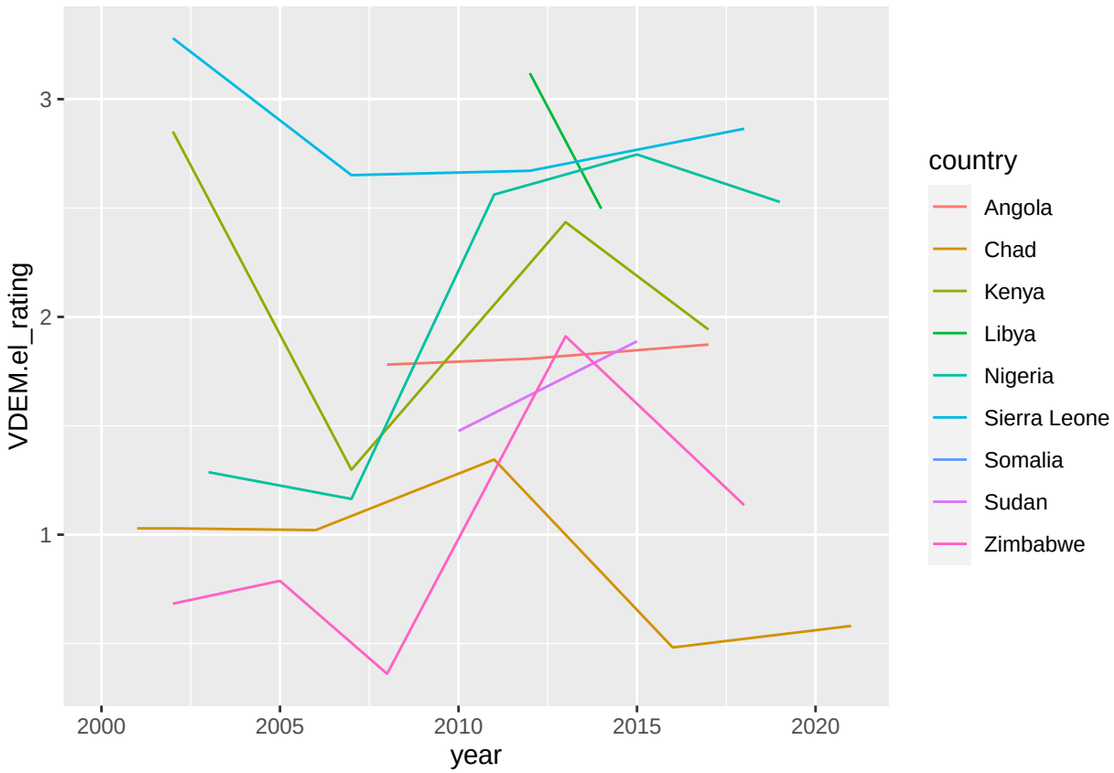
% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu % Date and time: Tue, May 17, 2022 - 16:58:11

Table 9:

	Dependent variable:	
	VDEM.el_rating	
	panel linear	*OLS*
	(1)	(2)
VDEM.EMB_cap	0.506***	0.600***
	(0.093)	(0.055)
Constant		0.913***
		(0.119)
Observations	283	283
R^2	0.115	0.300
Adjusted R^2	−0.081	0.298
Residual Std. Error		0.774 (df = 281)
F Statistic	29.912*** (df = 1; 231)	120.454*** (df = 1; 281)
Note:		*p<0.1; **p<0.05; ***p<0.01

Should make a linechart with year on x and el_rating on y. With each country as a line. Or maybe just IRI countries as a line. This might motivate a panel data chart.

```
iri_countries <- c("Kenya", "Chad", "Somalia", "Sudan", "Nigeria",
                   "Sierra Leone", "Congo", "Zimbabwe", "Angola", "Libya",
                   "Congo Kinshasa", "Congo (Kinshasa)", "Congo, Dem. Rep.")

ggplot(filter(merged_datasets3, country %in% iri_countries)) +
  geom_line(aes(year, VDEM.el_rating, color = country))
```




```
ggplot(merged_datasets3) +
  geom_line(aes(year, VDEM.el_rating, color = country))
```


```
#facet by income level and FH level
```

Lagged dependent variable

```
lagged.lm <- lm(VDEM.el_rating ~ VDEM.el_rating.lagged, merged_datasets3)
#stargazer(lagged.lm)

lagged.lm.incomelevel <- lm(VDEM.el_rating ~ VDEM.el_rating.lagged + income_level, merged_datasets3)
#stargazer(lagged.lm.incomelevel)

lagged.lm.income <- lm(VDEM.el_rating ~ VDEM.el_rating.lagged + WB.GDP_ppc, merged_datasets3)
lagged.lm.fh <- lm(VDEM.el_rating ~ VDEM.el_rating.lagged + FH.fh_total, merged_datasets3)
lagged.lm.fhlevel <- lm(VDEM.el_rating ~ VDEM.el_rating.lagged + fh_level, merged_datasets3)

stargazer(lagged.lm, lagged.lm.incomelevel, lagged.lm.income,lagged.lm.fh,lagged.lm.fhlevel)
```

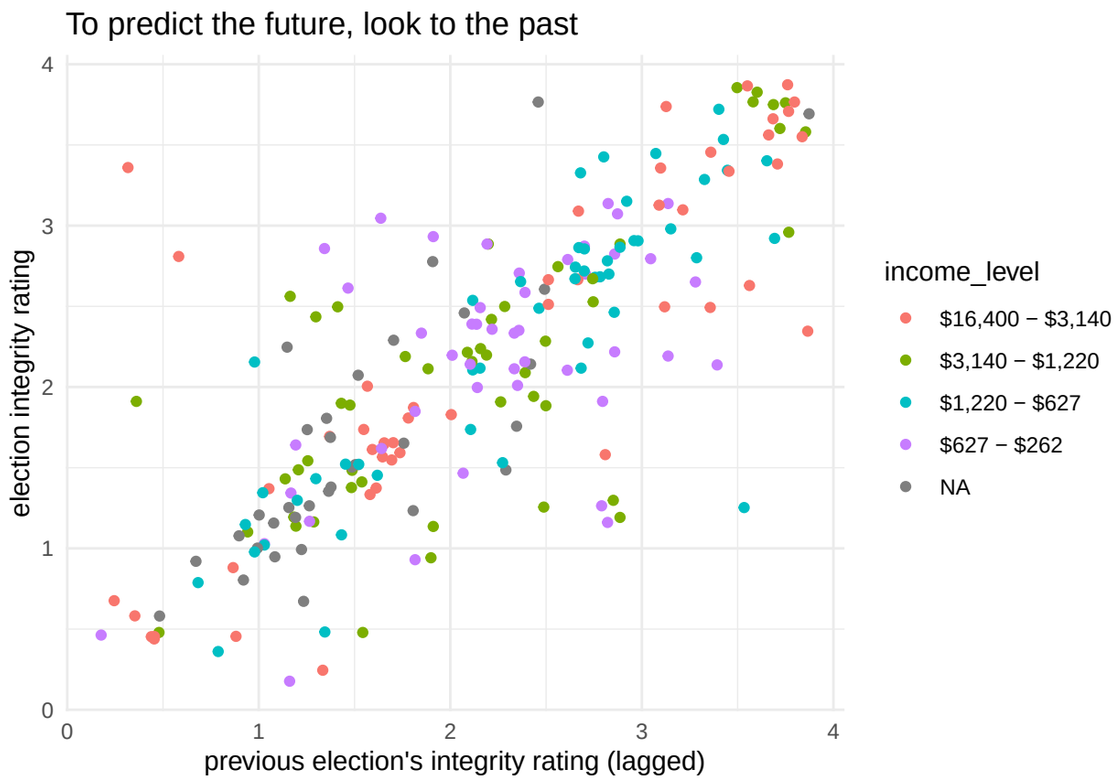
% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu % Date and time: Tue, May 17, 2022 - 16:58:16

```
p.income <- ggplot(merged_datasets3,
       aes(VDEM.el_rating.lagged, VDEM.el_rating, color = log(WB.GDP_ppc))) +
  geom_point() +
  theme_minimal() +
  labs(x = "previous election's integrity rating (lagged)",
       y = "election integrity rating") +
   #    title = "To predict the future, look to the past") +
    scale_color_continuous(name = "GDP per capita (log)", type = "viridis", labels=scales::dollar_format())


ggplot(merged_datasets3,
       aes(VDEM.el_rating.lagged, VDEM.el_rating, color = income_level)) +
  geom_point() +
  theme_minimal() +
  labs(x = "previous election's integrity rating (lagged)",
       y = "election integrity rating",
       title = "To predict the future, look to the past")
```

```
## Warning: Removed 81 rows containing missing values (geom_point).
```

Table 10:

	Dependent variable:				
	VDEM.el_rating				
	(1)	(2)	(3)	(4)	(5)
VDEM.el_rating.lagged	0.771***	0.754***	0.754***	0.415***	0.501***
	(0.041)	(0.046)	(0.046)	(0.063)	(0.059)
1,220		−0.048			
		(0.118)			
627		−0.041			
		(0.118)			
262		−0.056			
		(0.119)			
WB.GDP_ppc			0.00000		
			(0.00001)		
FH.fh_total				−0.135***	
				(0.020)	
fh_level2					−0.304***
					(0.112)
fh_level3					−0.567***
					(0.125)
fh_level4					−0.862***
					(0.148)
Constant	0.477***	0.555***	0.515***	2.426***	1.418***
	(0.094)	(0.131)	(0.115)	(0.294)	(0.185)
Observations	228	194	194	194	194
R^2	0.614	0.590	0.590	0.664	0.648
Adjusted R^2	0.613	0.582	0.585	0.661	0.640
Residual Std. Error	0.564 (df = 226)	0.585 (df = 189)	0.582 (df = 191)	0.520 (df = 191)	0.535 (df = 189)
F Statistic	359.839*** (df = 1; 226)	68.048*** (df = 4; 189)	28137.231*** (df = 2; 191)	188.909*** (df = 2; 191)	86.844*** (df = 4; 189)
Note:					*p<0.1; **p<0.05; ***p<0.01




```
p.fh <- ggplot(merged_datasets3,
       aes(VDEM.el_rating.lagged, VDEM.el_rating, color = FH.fh_total)) +
  geom_point() +
  theme_minimal() +
  labs(x = "previous election's integrity rating (lagged)",
       y = "election integrity rating") +
      # title = "To predict the future, look to the past") +
  scale_color_continuous(name = "Freedom House Scores \n(lower score = more free)", type = "viridis")

grid.arrange(p.income, p.fh, ncol = 2)
```

```
## Warning: Removed 81 rows containing missing values (geom_point).
```

```
## Warning: Removed 81 rows containing missing values (geom_point).
```