Individual-level analysis: simple models

Part of the final project for AQMSS II

Polikanov Stepan and Okisheva Vera

```
source(here::here("utilities", "check_packages.R"))
source(here::here("utilities", "functions.R"))


ep_raw_dep <- read_rds(here("data", "data_built", "ep_raw_dep.rds"))
data_country <- read_rds(here("data", "data_built", "data_country.rds"))


# Recode no data back to native NA
model_data <- ep_raw_dep |>
  mutate(across(c(sex, age_bin, time_to_vs.less_than_hour,
                  time_to_vs.less_than_hour, out_of_Russia_time,
                  result_trust_bin),
                ~ if_else(. %in% c("No Data", "Declined to answer"), NA, .)),
         vote = relevel(as.factor(vote), ref = "Putin"),
         sex = relevel(as.factor(sex), ref = "Male"),
         age_bin = relevel(as.factor(age_bin), ref = "25-44"),
         out_of_Russia_time = relevel(as.factor(out_of_Russia_time), ref = "Before annexation"),
         result_trust_bin = relevel(as.factor(result_trust_bin), ref = "Yes")) |>
  filter(!countryname_en %in% c("New Zealand", "Australia"))


coef_map_default <- c("(Intercept)", "sexFemale" = "Sex: Female",
                      "sexOther" = "Sex: Other",
                      "age_bin18-24" = "Age: 18-24 (ref 25-44)",
                      "age_bin45-64" = "Age: 45-65 (ref 25-44)",
                      "age_bin65+" = "Age: 65 + (ref 25-44)",
                      "time_to_vs.less_than_hourYes" =
                        "Took < 1 hour to get to the voting station",
                      "out_of_Russia_timeAfter invasion" =
                        "Moved after March 2022 (ref before 2014)",
                      "out_of_Russia_time2 - 5 years" =
                        paste("Moved after March 2019 but before",
                              "March 2022 (ref before 2014)"),
                      "out_of_Russia_timeAfter annexation" =
                        paste("Moved after March 2014 but before",
                              "March 2019 (ref before 2014)"),
                      "out_of_Russia_timeTourist (lives in Russia)" =
                        paste("Didn't move - tourist, lives",
                              "in Russia (ref before 2014)"),
                      "result_trust_binDon't know" =
                        "Trust in the result: Don't know (ref Yes)",
                      "result_trust_binNo" =
                        "Trust in the result: No (ref Yes)")
```

Table 1: Binary outcomes, linear models, naive approach

```
model_data |>
  pivot_longer(cols = c(vote_putin, vote_declined, vote_putin_declined,
                        vote_davankov, vote_spoiled, vote_opposition)) |>
  nest(data = -name) |>
  transmute(m1.naive = map(data,
                        ~ lm(value ~ sex + age_bin + time_to_vs.less_than_hour
                             + out_of_Russia_time + result_trust_bin,
                             data = .))) |>
  deframe() |>
  set_names(c("Vote Putin", "Decline to Answer", "Putin or Declined",
              "Vote Davankov", "Spoil the ballot", "Vote Davankov or spoil")) |>
  modelsummary(output = "kableExtra", vcov = "robust", stars = T,
               gof_map = c("nobs", "adj.r.squared", "bic", "logLik", "rmse"),
               coef_map = coef_map_default) |>
  kable_styling(latex_options = c("scale_down"))
```

	Vote Putin	Decline to Answer	Putin or Declined	Vote Davankov	Spoil the ballot	Vote Davankov or spoil
Sex: Female	0.010***	0.008***	0.018***	−0.042***	0.028***	−0.014***
	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)	(0.002)
Sex: Other	0.005	0.003	0.008	−0.087***	0.063**	−0.024+
	(0.012)	(0.013)	(0.012)	(0.023)	(0.022)	(0.014)
Age: 18-24 (ref 25-44)	−0.003	0.000	−0.003	0.005	−0.006	−0.001
	(0.002)	(0.002)	(0.003)	(0.006)	(0.006)	(0.004)
Age: 45-65 (ref 25-44)	0.044***	0.015***	0.059***	−0.154***	0.091***	−0.063***
	(0.004)	(0.004)	(0.003)	(0.006)	(0.005)	(0.004)
Age: 65 + (ref 25-44)	0.064***	0.029***	0.093***	−0.152***	0.053***	−0.099***
	(0.008)	(0.007)	(0.005)	(0.007)	(0.006)	(0.005)
Took < 1 hour to get to the voting station	0.001	0.013***	0.014***	−0.013***	−0.004	−0.017***
	(0.002)	(0.002)	(0.002)	(0.004)	(0.004)	(0.002)
Moved after March 2022 (ref before 2014)	−0.048***	−0.020***	−0.067***	0.169***	−0.088***	0.081***
	(0.003)	(0.003)	(0.003)	(0.006)	(0.005)	(0.004)
Moved after March 2019 but before March 2022 (ref before 2014)	−0.042***	−0.017***	−0.059***	0.113***	−0.051***	0.062***
	(0.004)	(0.004)	(0.004)	(0.007)	(0.006)	(0.004)
Moved after March 2014 but before March 2019 (ref before 2014)	−0.034***	−0.016***	−0.049***	0.050***	−0.001	0.049***
	(0.004)	(0.004)	(0.004)	(0.007)	(0.007)	(0.005)
Didn't move - tourist, lives in Russia (ref before 2014)	0.002	−0.015*	−0.013*	0.077***	−0.070***	0.007
	(0.007)	(0.007)	(0.006)	(0.009)	(0.007)	(0.007)
Trust in the result: Don't know (ref Yes)	−0.666***	0.031***	−0.634***	0.436***	0.136***	0.572***
	(0.008)	(0.009)	(0.010)	(0.011)	(0.007)	(0.011)
Trust in the result: No (ref Yes)	−0.731***	−0.081***	−0.812***	0.458***	0.328***	0.786***
	(0.005)	(0.004)	(0.005)	(0.005)	(0.004)	(0.005)
Num.Obs.	54 111	54 111	54 111	54 111	54 111	54 111
R2 Adj.	0.726	0.047	0.763	0.291	0.082	0.668
BIC	−24 139.2	−15 371.5	−20 960.5	59 423.4	50 050.2	2406.8
Log.Lik.	12 145.880	7762.046	10 556.535	−29 635.417	−24 948.803	−1127.086
RMSE	0.19	0.21	0.20	0.42	0.38	0.25

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

Table 2: Multinomial regression, naive approach

	Naive approach					
	Davankov	Declined to answer	Haritonov	Slutsky	Spoiled ballot	Tore up/took
Sex: Female	−0.565***	−0.070	−0.590***	−0.471***	−0.339***	−0.691***
	(0.056)	(0.052)	(0.084)	(0.102)	(0.059)	(0.132)
Sex: Other	−0.503	−0.045	−0.130	0.326	0.022	0.400
	(0.314)	(0.267)	(0.466)	(0.489)	(0.324)	(0.588)
Age: 18-24 (ref 25-44)	0.114	0.106	0.069	0.384*	0.087	0.423*
	(0.105)	(0.111)	(0.147)	(0.161)	(0.109)	(0.203)
Age: 45-65 (ref 25-44)	−1.308***	−0.213***	−0.769***	−0.733***	−0.452***	−0.171
	(0.069)	(0.061)	(0.119)	(0.147)	(0.073)	(0.179)
Age: 65 + (ref 25-44)	−1.870***	−0.091	−0.507**	−0.800**	−0.871***	−0.414
	(0.118)	(0.076)	(0.183)	(0.248)	(0.127)	(0.383)
Took < 1 hour to get to the voting station	−0.326***	0.163**	−0.236**	−0.169	−0.323***	−0.208
	(0.062)	(0.059)	(0.091)	(0.111)	(0.064)	(0.139)
Moved after March 2022 (ref before 2014)	1.877***	0.623***	0.895***	1.086***	1.120***	1.083***
	(0.074)	(0.071)	(0.115)	(0.146)	(0.078)	(0.193)
Moved after March 2019 but before March 2022 (ref before 2014)	1.215***	0.209*	0.638***	0.814***	0.689***	0.975***
	(0.088)	(0.087)	(0.133)	(0.166)	(0.092)	(0.212)
Moved after March 2014 but before March 2019 (ref before 2014)	0.759***	−0.012	0.265+	0.653***	0.532***	0.708**
	(0.096)	(0.093)	(0.151)	(0.179)	(0.100)	(0.234)
Didn't move - tourist, lives in Russia (ref before 2014)	0.332**	−0.244*	−0.069	0.517*	−0.524***	−0.206
	(0.113)	(0.099)	(0.201)	(0.216)	(0.130)	(0.395)
Trust in the result: Don't know (ref Yes)	4.100***	2.287***	3.350***	3.594***	3.890***	2.617***
	(0.105)	(0.110)	(0.186)	(0.192)	(0.156)	(0.667)
Trust in the result: No (ref Yes)	7.153***	3.648***	5.779***	5.259***	8.187***	7.097***
	(0.099)	(0.100)	(0.145)	(0.161)	(0.130)	(0.341)
Num.Obs.	54 111					
R2 Adj.	0.533					
BIC	89 346.1					
RMSE	0.25					

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

Multinomial

```
m2.naive <- multinom(vote ~ sex + age_bin + time_to_vs.less_than_hour
                     + out_of_Russia_time + result_trust_bin,
                     data = model_data, model = F)
```

```
m2.fe <- multinom(vote ~ sex + age_bin + time_to_vs.less_than_hour
                  + out_of_Russia_time + result_trust_bin
                  + as.factor(voting_station),
                  data = model_data, model = F)
```

```
modelsummary(list("Naive approach" = m2.naive),
             output = "kableExtra",
             stars = T, shape = term ~ response,
             coef_map = coef_map_default,
             gof_map = c("nobs", "adj.r.squared", "bic", "rmse")) |>
  kable_styling(latex_options = c("scale_down"))
```

```
gc()
```

```
            used  (Mb) gc trigger  (Mb) max used  (Mb)
```

Table 3: Multinomial regression, fixed effects

	Fixed Effects					
	Davankov	Declined to answer	Haritonov	Slutsky	Spoiled ballot	Tore up/took
Sex: Female	−0.452***	0.092+	−0.403***	−0.321**	−0.248***	−0.644***
	(0.059)	(0.054)	(0.090)	(0.113)	(0.062)	(0.147)
Sex: Other	−0.679*	−0.232	−0.302	0.196	−0.174	0.257
	(0.318)	(0.274)	(0.489)	(0.517)	(0.328)	(0.617)
Age: 18-24 (ref 25-44)	0.072	0.088	−0.115	0.324+	0.001	0.466*
	(0.112)	(0.117)	(0.164)	(0.180)	(0.116)	(0.227)
Age: 45-65 (ref 25-44)	−1.289***	−0.287***	−0.735***	−0.715***	−0.439***	0.004
	(0.073)	(0.063)	(0.127)	(0.163)	(0.077)	(0.195)
Age: 65 + (ref 25-44)	−1.757***	−0.387***	−0.350+	−0.763**	−0.811***	−0.399
	(0.120)	(0.082)	(0.197)	(0.292)	(0.133)	(0.465)
Took < 1 hour to get to the voting station	−0.030	0.273***	0.113	0.075	0.001	0.050
	(0.071)	(0.066)	(0.107)	(0.134)	(0.073)	(0.172)
Moved after March 2022 (ref before 2014)	2.183***	0.796***	1.319***	1.300***	1.647***	1.704***
	(0.087)	(0.083)	(0.137)	(0.179)	(0.092)	(0.240)
Moved after March 2019 but before March 2022 (ref before 2014)	1.399***	0.402***	0.869***	0.974***	0.929***	1.316***
	(0.096)	(0.093)	(0.146)	(0.187)	(0.100)	(0.248)
Moved after March 2014 but before March 2019 (ref before 2014)	0.854***	0.154	0.378*	0.846***	0.619***	1.030***
	(0.101)	(0.096)	(0.161)	(0.195)	(0.106)	(0.260)
Didn't move - tourist, lives in Russia (ref before 2014)	0.786***	0.231*	0.373+	0.949***	0.197	0.609
	(0.126)	(0.109)	(0.225)	(0.254)	(0.142)	(0.467)
Trust in the result: Don't know (ref Yes)	3.918***	2.318***	3.527***	3.927***	3.713***	−3.139***
	(0.109)	(0.115)	(0.198)	(0.225)	(0.159)	(0.005)
Trust in the result: No (ref Yes)	7.630***	4.343***	6.554***	6.311***	8.689***	11.848***
	(0.137)	(0.138)	(0.184)	(0.215)	(0.160)	(2.603)
Num.Obs.	54 111					
R2 Adj.	0.542					
BIC	91 620.5					
RMSE	0.25					

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

```
Ncells  5151648 275.2    8854521 472.9  7003970 374.1
Vcells 13252724 101.2   49532909 378.0 49532906 378.0
```

```
modelsummary(list("Fixed Effects" = m2.fe),
             output = "kableExtra",
             stars = T, shape = term ~ response,
             coef_map = coef_map_default,
             gof_map = c("nobs", "adj.r.squared", "bic", "rmse")) |>
  kable_styling(latex_options = c("scale_down"))
```

```
comparisons <- logits(answer = dichotomy(answer = c("Davankov",
                                                    "Spoiled ballot",
                                                    "Slutsky",
                                                    "Haritonov",
                                                    "Putin"),
                                         "Declined to answer"),
                      not_putin = dichotomy(opposition = c("Davankov",
                                                           "Spoiled ballot",
                                                           "Slutsky",
                                                           "Haritonov"),
                                            "Putin"),
                      opposition = dichotomy(
                        systemic = c("Slutsky", "Haritonov"),
                        nonsystemic = c("Davankov", "Spoiled ballot")),
                      nonsystemic = c("Spoiled ballot", "Davankov"),
                      systemic = c("Haritonov", "Slutsky"))
```

```r
m3.nested.naive <- nestedLogit(vote ~ sex + age_bin + time_to_vs.less_than_hour
                               + out_of_Russia_time + result_trust_bin,
                               dichotomies = comparisons,
                               data = model_data,
                               subset = model_data$vote != "Tore up/took"
                                        | !is.na(model_data$vote))

m3.nested.fe <- nestedLogit(vote ~ sex + age_bin + time_to_vs.less_than_hour
                            + out_of_Russia_time + result_trust_bin + as.factor(voting_station),
                            dichotomies = comparisons,
                            data = model_data,
                            subset = model_data$vote != "Tore up/took"
                            | !is.na(model_data$vote))

save(list = c("m3.nested.naive", "m3.nested.fe"),
     file = here("scripts", "models", "nl_fe.RData"))
```

Table 4

```
resizebox.stargazer(models(m3.nested.naive),
          title = "Nested Logit models", header = F,
          dep.var.labels = c("Don't answer vs answer", "Putin vs everyone",
                             "Non-systemic vs systemic opposition",
                             "Spoiled vs Davankov", "Slutsky vs Haritonov"),
          covariate.labels = c(
            "Sex: Female", "Sex: Other", "Age: 18-24 (ref 25-44)",
            "Age: 45-65 (ref 25-44)", "Age: 65 + (ref 25-44)",
            "Took < 1 hour to get to the voting station",
            "Moved after March 2022 (ref before 2014)",
            "Moved after March 2019 but before March 2022 (ref before 2014)",
            "Moved after March 2014 but before March 2019 (ref before 2014)",
            "Didn't move - tourist, lives in Russia (ref before 2014)",
            "Trust in the result: Don't know (ref Yes)",
            "Trust in the result: No (ref Yes)", "Intercept"),
          tab.height = "\\textheight", tab.width= "\\textwidth"
          )
```

[!htbp]

Table 5Nested Logit models

	Dependent variable:				
	Don't answer vs answer	Putin vs everyone	Non-systemic vs systemic opposition	Spoiled vs Davankov	Slutsky vs Haritonov
	(1)	(2)	(3)	(4)	(5)
Sex: Female	0.192***	0.510***	0.022	−0.220***	0.155
	(0.043)	(0.069)	(0.056)	(0.024)	(0.115)
Sex: Other	0.088	−0.138	−0.383	−0.517***	0.423
	(0.234)	(0.372)	(0.298)	(0.146)	(0.573)
Age: 18-24 (ref 25-44)	0.004	−0.132	−0.079	0.025	0.327*
	(0.080)	(0.118)	(0.085)	(0.036)	(0.170)
Age: 45-65 (ref 25-44)	0.249***	1.336***	−0.308***	−0.836***	0.007
	(0.057)	(0.083)	(0.087)	(0.039)	(0.179)
Age: 65 + (ref 25-44)	0.339***	1.542***	−0.939***	−0.959***	−0.424
	(0.074)	(0.127)	(0.154)	(0.106)	(0.305)
Took < 1 hour to get to the voting station	0.349***	0.368***	−0.107*	−0.001	0.034
	(0.049)	(0.075)	(0.059)	(0.024)	(0.122)
Moved after March 2022 (ref before 2014)	−0.294***	−1.413***	0.401***	0.513***	0.148
	(0.073)	(0.103)	(0.090)	(0.041)	(0.187)
Moved after March 2019 but before March 2022 (ref before 2014)	−0.227***	−0.816***	0.292***	0.217***	0.387*
	(0.078)	(0.116)	(0.102)	(0.045)	(0.208)
Moved after March 2014 but before March 2019 (ref before 2014)	−0.412***	−1.991***	0.741***	0.742***	0.163
	(0.062)	(0.087)	(0.080)	(0.036)	(0.166)
Didn't move - tourist, lives in Russia (ref before 2014)	−0.200**	−0.530***	−0.001	0.858***	0.542**
	(0.090)	(0.130)	(0.141)	(0.084)	(0.274)
Trust in the result: Don't know (ref Yes)	0.438***	−4.040***	0.708***	0.403***	0.253
	(0.080)	(0.105)	(0.127)	(0.140)	(0.239)
Trust in the result: No (ref Yes)	−1.381***	−7.365***	1.913***	−0.809***	−0.548***
	(0.057)	(0.099)	(0.094)	(0.106)	(0.183)
Intercept	−2.404***	1.881***	1.276***	1.466***	−0.466**
	(0.072)	(0.098)	(0.119)	(0.111)	(0.229)
Observations	53,824	51,202	42,363	40,946	1,417
Log Likelihood	−9,435.706	−3,530.187	−5,869.120	−22,756.890	−911.162
Akaike Inf. Crit.	18,897.410	7,086.374	11,764.240	45,539.780	1,848.323

Note: *p<0.1; **p<0.05; ***p<0.01

```
resizebox.stargazer(models(m3.nested.fe),
          title = "Nested Logit models, fixed effects", header = F,
          dep.var.labels = c("Don't answer vs answer", "Putin vs everyone",
                             "Non-systemic vs systemic opposition",
                             "Spoiled vs Davankov", "Slutsky vs Haritonov"),
          covariate.labels = c(
            "Sex: Female", "Sex: Other", "Age: 18-24 (ref 25-44)",
            "Age: 45-65 (ref 25-44)", "Age: 65 + (ref 25-44)",
            "Took < 1 hour to get to the voting station",
            "Moved after March 2022 (ref before 2014)",
            "Moved after March 2019 but before March 2022 (ref before 2014)",
            "Moved after March 2014 but before March 2019 (ref before 2014)",
            "Didn't move - tourist, lives in Russia (ref before 2014)",
            "Trust in the result: Don't know (ref Yes)",
            "Trust in the result: No (ref Yes)", "Intercept"),
          omit = "as.factor",
          tab.height = "\\textheight", tab.width= "\\textwidth"
```