Replication script for manuscript 'Subnational political actors and socialization in European institutions'

For peer review

01/12/2023

This document is prepared to present the analyses in the paper "Subnational political actors and socialization in European institutions: An analysis of views on multilevel governance" and to facilitate replication.

The necessary packages are loaded as follows:

```
library(tidyverse) #for general data pipelines
library(simputation) #for imputation of missing data
library(skimr) #for descriptives
library(bestNormalize) #for normalizing numeric variables
library(lmtest) #for robust standard errors
library(sandwich) #for robust standard errors
```

We load the data from the csv file and subset for relevant observations and variables:

```
df <- read_csv("subnationalViews_dataset.csv",
               col_select = c(surveyId, lang, country, sneType,
                              mlgS1, mlgS2, mlgS3, mlgS4, mlgS5, mlgC1, mlgC2,
                              euCompSocpol, euCompEc, euCompInAf, euCompForAf,
                              euCompOv, snaCompSocpol, snaCompEc, snaCompInAf,
                              snaCompForAf, snaCompOv, snaCompC1, snaCompC2,
                              effectiveIntOrg, effectiveDirRel, effectiveNatDel,
                              euFundImp.bin, represent, represent.bin, repDur,
                              norepMotiv, age, gender, eduLev, polLeftRight,
                              polIdLoc, polIdNat, polIdEu, polIdCul, polIdIdeo)) %>%
  filter(country %in% c("Belgium", "France", "Germany", "Ireland", "Italy",
                        "Netherlands", "Norway", "Spain", "Sweden", "Switzerland",
                        "United Kingdom"))
```

We use KNN imputation to deal with missing data:

```
set.seed(2023)
df_imp <- df %>%
  impute_knn( . -surveyId -lang -country -countryPop -sneType -represent
              -represent.bin -norepMotiv -repDur ~
                . -surveyId -lang -country -countryPop -sneType -represent
              -represent.bin -norepMotiv -repDur |
                country + sneType)
```

One case cannot be completed for variables age and political orientation; we drop it

```
df_imp <- df_imp %>% filter(!is.na(age) & !is.na(polLeftRight))
```

The number of observations per country is unbalanced; for this reason we calculate weights to treat each country evenly, which will be used in further analyses. However, later stages of analyses use different subsets of data; for this reason, we define a function for the calculation of weights:

```
weight_calc <- function(df) {
  wt_tbl <- df %>%
    group_by(country) %>%
    summarise(freq_ctr = n()) %>%
    ungroup() %>%
    mutate(perc_ctr = freq_ctr / sum(freq_ctr),
           weight = mean(perc_ctr) / perc_ctr)
  df <- df %>%
    left_join(wt_tbl, by = "country", suffix = c("_old", "")) %>%
    select(-freq_ctr, -perc_ctr)
  return(df)
}

df_imp <- weight_calc(df_imp)
```

We correct variable types and category levels as necessary:

```
df_imp <- df_imp %>%
  mutate(surveyId = as.character(surveyId),
         across(c(lang, country, sneType, effectiveIntOrg:euFundImp.bin, represent.bin,
                  gender, polIdLoc:polIdIdeo), as.factor),
         across(c(mlgS1:snaCompC2, age), as.numeric),
         represent = factor(represent, levels=c("NONE", "COR", "CLRAE", "OTHER")),
         norepMotiv = factor(norepMotiv, levels=as.character(1:5), ordered=TRUE),
         eduLev = factor(eduLev, levels=as.character(1:3), ordered=TRUE),
         polLeftRight = factor(polLeftRight, levels=as.character(1:9), ordered=TRUE))
```

Principal component analysis (PCA) to reduce answers to fewer composite index variables:

```
pca_mlg <- prcomp(df_imp[5:11], scale.=TRUE)
pca_eu <- prcomp(df_imp[12:16], scale.=TRUE)
pca_sna <- prcomp(df_imp[17:23], scale.=TRUE)

#summary(pca_mlg) ; print(pca_mlg)
#summary(pca_eu) ; print(pca_eu)
#summary(pca_sna) ; print(pca_sna)
```

We calculate the index variables with weights from PCA:

```
df_imp <- df_imp %>%
  rowwise() %>%
  mutate(mlgInd = weighted.mean(c(mlgS1, mlgS2, mlgS3, mlgS4, mlgS5, mlgC1, mlgC2),
                                pca_mlg$rotation[,1]),
         euInd = weighted.mean(c(euCompSocpol, euCompEc, euCompInAf, euCompForAf, euCompOv),
                               pca_eu$rotation[,1]),
```

```
        snaInd = weighted.mean(c(snaCompSocpol, snaCompEc, snaCompInAf, snaCompForAf,
                                 snaCompOv, snaCompC1, snaCompC2),
                               pca_sna$rotation[,1])) %>%
  ungroup()
```

We simplify political ideology into three groups:

```
df_imp <- df_imp %>%
  mutate(LRC = case_when(
           as.numeric(polLeftRight) <=3 ~ "Left",
           as.numeric(polLeftRight) >3 & as.numeric(polLeftRight) <=6 ~ "Centre",
           as.numeric(polLeftRight) >6 ~ "Right"
         ) %>% factor(levels = c("Centre", "Left", "Right")))
```

Descriptive statistics:

```
skim(df_imp)
```

Table 1: Data summary

Name	df_imp
Number of rows	648
Number of columns	45
Column type frequency:	
character	1
factor	19
numeric	25
Group variables	None

Variable type: character

skim_variable	n_missing	complete_rate	min	max	empty	n_unique	whitespace
surveyId	0	1	2	4	0	648	0

Variable type: factor

skim_variable	n_missing	complete_rate	ordered	n_unique	top_counts
lang	0	1.00	FALSE	5	en: 266, de: 185, fr: 125, it: 47
country	0	1.00	FALSE	11	Swi: 150, Uni: 95, Ger: 92, Ita: 47
sneType	0	1.00	FALSE	4	Loc: 321, Reg: 213, Reg: 112, Oth: 2
effectiveIntOrg	0	1.00	FALSE	2	1: 370, 0: 278
effectiveDirRel	0	1.00	FALSE	2	1: 579, 0: 69
effectiveNatDel	0	1.00	FALSE	2	1: 402, 0: 246
euFundImp.bin	0	1.00	FALSE	2	1: 491, 0: 157

skim_variable	n_missing	complete_rate	ordered	n_unique	top_counts
represent	0	1.00	FALSE	4	NON: 512, OTH: 75, COR: 39, CLR: 22
represent.bin	0	1.00	FALSE	2	0: 512, 1: 136
norepMotiv	150	0.77	TRUE	5	4: 172, 5: 154, 3: 62, 2: 56
gender	0	1.00	FALSE	3	Mal: 452, Fem: 193, Oth: 3
eduLev	0	1.00	TRUE	3	3: 325, 2: 257, 1: 66
polLeftRight	0	1.00	TRUE	9	3: 132, 2: 110, 7: 105, 4: 83
polIdLoc	0	1.00	FALSE	2	1: 601, 0: 47
polIdNat	0	1.00	FALSE	2	0: 402, 1: 246
polIdEu	0	1.00	FALSE	2	0: 498, 1: 150
polIdCul	0	1.00	FALSE	2	0: 582, 1: 66
polIdIdeo	0	1.00	FALSE	2	1: 517, 0: 131
LRC	0	1.00	FALSE	3	Lef: 272, Cen: 231, Rig: 145

Variable type: numeric

skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100	hist
mlgS1	0	1.0	3.79	1.26	1.00	3.00	4.00	5.00	5.00	
mlgS2	0	1.0	4.20	1.10	1.00	4.00	5.00	5.00	5.00	
mlgS3	0	1.0	3.79	1.30	1.00	3.00	4.00	5.00	5.00	
mlgS4	0	1.0	3.65	1.35	1.00	3.00	4.00	5.00	5.00	
mlgS5	0	1.0	3.39	1.39	1.00	2.00	4.00	5.00	5.00	
mlgC1	0	1.0	3.72	1.10	1.00	3.00	4.00	5.00	5.00	
mlgC2	0	1.0	4.06	1.06	1.00	4.00	4.00	5.00	5.00	
euCompSocpol	0	1.0	2.92	1.30	1.00	2.00	3.00	4.00	5.00	
euCompEc	0	1.0	3.76	1.17	1.00	3.00	4.00	5.00	5.00	
euCompInAf	0	1.0	3.46	1.29	1.00	2.00	4.00	4.00	5.00	
euCompForAf	0	1.0	3.98	1.18	1.00	4.00	4.00	5.00	5.00	
euCompOv	0	1.0	3.48	1.04	1.00	3.00	4.00	4.00	5.00	
snaCompSocpol	0	1.0	3.84	1.07	1.00	3.00	4.00	5.00	5.00	
snaCompEc	0	1.0	3.32	1.18	1.00	2.00	4.00	4.00	5.00	
snaCompInAf	0	1.0	2.92	1.25	1.00	2.00	3.00	4.00	5.00	
snaCompForAf	0	1.0	2.47	1.26	1.00	1.00	2.00	3.00	5.00	
snaCompOv	0	1.0	3.52	0.94	1.00	3.00	4.00	4.00	5.00	
snaCompC1	0	1.0	3.82	1.25	1.00	3.00	4.00	5.00	5.00	
snaCompC2	0	1.0	4.03	1.02	1.00	4.00	4.00	5.00	5.00	
repDur	516	0.2	3.77	1.35	1.00	3.00	4.00	5.00	5.00	
age	0	1.0	49.82	12.47	21.00	40.00	52.00	59.00	83.00	
weight	0	1.0	1.00	0.56	0.39	0.62	0.64	1.37	2.36	
mlgInd	0	1.0	3.78	0.88	1.00	3.33	3.93	4.50	5.00	
euInd	0	1.0	3.53	0.94	1.00	3.01	3.80	4.20	5.00	
snaInd	0	1.0	3.30	0.80	1.06	2.75	3.31	3.92	5.00	

```
df_imp %>% summarise(across(where(is.factor), ~ sum(.x == 1)/n()))
```

```
## # A tibble: 1 x 19
##    lang country sneType effectiveIntOrg effectiveDirRel effectiveNatDel
##   <dbl>   <dbl>   <dbl>           <dbl>           <dbl>           <dbl>
## 1     0       0       0           0.571           0.894           0.620
## # i 13 more variables: euFundImp.bin <dbl>, represent <dbl>,
```

```
## #   represent.bin <dbl>, norepMotiv <dbl>, gender <dbl>, eduLev <dbl>,
## #   polLeftRight <dbl>, polIdLoc <dbl>, polIdNat <dbl>, polIdEu <dbl>,
## #   polIdCul <dbl>, polIdIdeo <dbl>, LRC <dbl>
```

```
df_imp %>% summarise(left = sum(LRC=="Left")/n(),
                     right = sum(LRC=="Right")/n(),
                     centre = sum(LRC=="Centre")/(n()),
                     female = sum(gender=="Female")/n(),
                     male = sum(gender=="Male")/n(),
                     edu1 = sum(eduLev==1)/n(),
                     edu2 = sum(eduLev==2)/n(),
                     edu3 = sum(eduLev==3)/n(),
                     sneLocal = sum(sneType=="Local")/n(),
                     sneRegAd = sum(sneType=="RegionAd")/n(),
                     sneRegLeg = sum(sneType=="RegionLeg")/n())
```

```
## # A tibble: 1 x 11
##    left right centre female  male  edu1  edu2  edu3 sneLocal sneRegAd sneRegLeg
##   <dbl> <dbl>  <dbl>  <dbl> <dbl> <dbl> <dbl> <dbl>    <dbl>    <dbl>     <dbl>
## 1 0.420 0.224  0.356  0.298 0.698 0.102 0.397 0.502    0.495    0.173     0.329
```

```
summary(as.numeric(df_imp$norepMotiv)) ; sd(as.numeric(df_imp$norepMotiv), na.rm=TRUE)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   1.000   3.000   4.000   3.635   5.000   5.000     150
```

```
## [1] 1.315047
```

We normalise the index variables with standardised values for regression analyses:

```
df_imp <- df_imp %>%
  mutate(mlgInd.norm = orderNorm(mlgInd)[[1]],
         euInd.norm = orderNorm(euInd)[[1]],
         snaInd.norm = orderNorm(snaInd)[[1]])
```

Fitting the base model: the first output is with normal standard errors, the second with robust:

```
mainFormula <- as.formula("mlgInd.norm ~ country + represent.bin +
                          euInd.norm + snaInd.norm + effectiveIntOrg +
                          effectiveDirRel + euFundImp.bin + polIdEu +
                          polIdCul + polIdIdeo + LRC + age + gender + eduLev")

model_base <- lm(mainFormula, data = df_imp, weights = weight)

summary(model_base)
```

```
##
## Call:
## lm(formula = mainFormula, data = df_imp, weights = weight)
##
## Weighted Residuals:
```

```
##     Min      1Q  Median      3Q     Max
## -4.0902 -0.4224 -0.0053  0.3967  2.3026
##
## Coefficients:
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)           -1.037208   0.225812  -4.593 5.29e-06 ***
## countryFrance          0.391100   0.145597   2.686  0.00742 **
## countryGermany         0.337896   0.149294   2.263  0.02396 *
## countryIreland         0.223532   0.146752   1.523  0.12822
## countryItaly           0.348438   0.148974   2.339  0.01966 *
## countryNetherlands     0.157849   0.145156   1.087  0.27726
## countryNorway          0.154860   0.147494   1.050  0.29416
## countrySpain           0.041833   0.146536   0.285  0.77537
## countrySweden          0.225884   0.148068   1.526  0.12763
## countrySwitzerland     0.151869   0.149244   1.018  0.30927
## countryUnited Kingdom  0.298294   0.147024   2.029  0.04290 *
## represent.bin1         0.388753   0.077285   5.030 6.43e-07 ***
## euInd.norm             0.192431   0.036568   5.262 1.96e-07 ***
## snaInd.norm            0.302007   0.032045   9.425  < 2e-16 ***
## effectiveIntOrg1       0.365185   0.072548   5.034 6.31e-07 ***
## effectiveDirRel1       0.302144   0.110761   2.728  0.00655 **
## euFundImp.bin1         0.390570   0.092315   4.231 2.68e-05 ***
## polIdEu1               0.003339   0.077045   0.043  0.96545
## polIdCul1              0.411619   0.102221   4.027 6.36e-05 ***
## polIdIdeo1            -0.156259   0.079245  -1.972  0.04907 *
## LRCLeft                0.063232   0.071170   0.888  0.37464
## LRCRight              -0.148681   0.085464  -1.740  0.08241 .
## age                    0.001211   0.002617   0.463  0.64358
## genderMale            -0.043369   0.068683  -0.631  0.52799
## genderOther            0.122207   0.406193   0.301  0.76362
## eduLev.L               0.058400   0.076572   0.763  0.44595
## eduLev.Q               0.035360   0.060813   0.581  0.56114
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7699 on 621 degrees of freedom
## Multiple R-squared:  0.4362, Adjusted R-squared:  0.4126
## F-statistic: 18.48 on 26 and 621 DF,  p-value: < 2.2e-16
```

```
coeftest(model_base, vcov=vcovHC(model_base, type="HC2", cluster = ~country))
```

```
##
## t test of coefficients:
##
##                          Estimate Std. Error t value  Pr(>|t|)
## (Intercept)           -1.0372082  0.2481745 -4.1793 3.343e-05 ***
## countryFrance          0.3910995  0.1693978  2.3088 0.0212842 *
## countryGermany         0.3378960  0.1512636  2.2338 0.0258500 *
## countryIreland         0.2235317  0.1658031  1.3482 0.1780935
## countryItaly           0.3484384  0.1584218  2.1994 0.0282148 *
## countryNetherlands     0.1578491  0.1708081  0.9241 0.3557771
## countryNorway          0.1548596  0.1642258  0.9430 0.3460643
## countrySpain           0.0418333  0.2718482  0.1539 0.8777505
## countrySweden          0.2258839  0.1714755  1.3173 0.1882254
```

```
## countrySwitzerland     0.1518693  0.1449542  1.0477 0.2951819
## countryUnited Kingdom  0.2982942  0.1518395  1.9645 0.0499135 *
## represent.bin1         0.3887534  0.0938154  4.1438 3.890e-05 ***
## euInd.norm             0.1924312  0.0538866  3.5710 0.0003831 ***
## snaInd.norm            0.3020068  0.0461255  6.5475 1.227e-10 ***
## effectiveIntOrg1       0.3651849  0.0780149  4.6810 3.508e-06 ***
## effectiveDirRel1       0.3021436  0.1145061  2.6387 0.0085320 **
## euFundImp.bin1         0.3905702  0.1161860  3.3616 0.0008225 ***
## polIdEu1               0.0033390  0.0923322  0.0362 0.9711640
## polIdCul1              0.4116190  0.1224867  3.3605 0.0008257 ***
## polIdIdeo1            -0.1562591  0.0913681 -1.7102 0.0877256 .
## LRCLeft                0.0632321  0.0820574  0.7706 0.4412469
## LRCRight              -0.1486807  0.0977765 -1.5206 0.1288646
## age                    0.0012113  0.0028265  0.4285 0.6683996
## genderMale            -0.0433687  0.0797833 -0.5436 0.5869249
## genderOther            0.1222075  0.1431383  0.8538 0.3935606
## eduLev.L               0.0583996  0.0858370  0.6804 0.4965334
## eduLev.Q               0.0353604  0.0732237  0.4829 0.6293304
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Diagnostics: heteroskedasticity, outliers, multicolinearity

```
bptest(model_base)
```

```
##
##  studentized Breusch-Pagan test
##
## data:  model_base
## BP = 81.278, df = 26, p-value = 1.33e-07
```

```
car::outlierTest(model_base)
```

```
##      rstudent unadjusted p-value Bonferroni p
## 407 -5.807912         1.0111e-08   6.5517e-06
## 565 -5.207786         2.6038e-07   1.6873e-04
## 430 -4.350279         1.5897e-05   1.0301e-02
## 167 -4.333073         1.7152e-05   1.1114e-02
```

```
car::vif(model_base)
```

```
##                     GVIF Df GVIF^(1/(2*Df))
## country         2.950261 10        1.055585
## represent.bin   1.181606  1        1.087017
## euInd.norm      1.378072  1        1.173913
## snaInd.norm     1.239404  1        1.113285
## effectiveIntOrg 1.320986  1        1.149342
## effectiveDirRel 1.169177  1        1.081285
## euFundImp.bin   1.390258  1        1.179092
## polIdEu         1.264432  1        1.124470
## polIdCul        1.117068  1        1.056915
## polIdIdeo       1.180467  1        1.086493
```

```
## LRC             1.262051  2        1.059911
## age             1.160894  1        1.077448
## gender          1.176940  2        1.041570
## eduLev          1.253022  2        1.058010
```

We increase the robustness of the model by dropping outliers that cause heteroskedasticity. Dropping the third outlier does not reduce heteroskedasticity, keeping it does not lead to Type I error (case contradicting posited relationship); it is kept in the sample. Dropping other three outliers reduce heteroskedasticity without leading to Type I error (cases confirming posited relationship); they are dropped.

The heteroskedasticity in the robust model is not statistically significant and not visually obvious. The output indeed shows that the coefficient estimate of the main predictor is lower than the base model.

```
outlier_names <- names(car::outlierTest(model_base)[[1]])
df_imp %>% filter(row.names(df_imp) %in% outlier_names)
```

```
## # A tibble: 4 x 48
##   surveyId lang  country sneType   mlgS1 mlgS2 mlgS3 mlgS4 mlgS5 mlgC1 mlgC2
##   <chr>    <fct> <fct>   <fct>     <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 490      es    Spain   RegionAd      4     1     1     1     1     4     5
## 2 1074     es    Spain   RegionLeg     1     1     1     1     1     1     1
## 3 1113     fr    Belgium RegionAd      1     1     1     1     2     5     5
## 4 1500     es    Spain   RegionLeg     1     1     1     1     1     5     5
## # i 37 more variables: euCompSocpol <dbl>, euCompEc <dbl>, euCompInAf <dbl>,
## #   euCompForAf <dbl>, euCompOv <dbl>, snaCompSocpol <dbl>, snaCompEc <dbl>,
## #   snaCompInAf <dbl>, snaCompForAf <dbl>, snaCompOv <dbl>, snaCompC1 <dbl>,
## #   snaCompC2 <dbl>, effectiveIntOrg <fct>, effectiveDirRel <fct>,
## #   effectiveNatDel <fct>, euFundImp.bin <fct>, represent <fct>,
## #   represent.bin <fct>, repDur <dbl>, norepMotiv <ord>, age <dbl>,
## #   gender <fct>, eduLev <ord>, polLeftRight <ord>, polIdLoc <fct>, ...
```

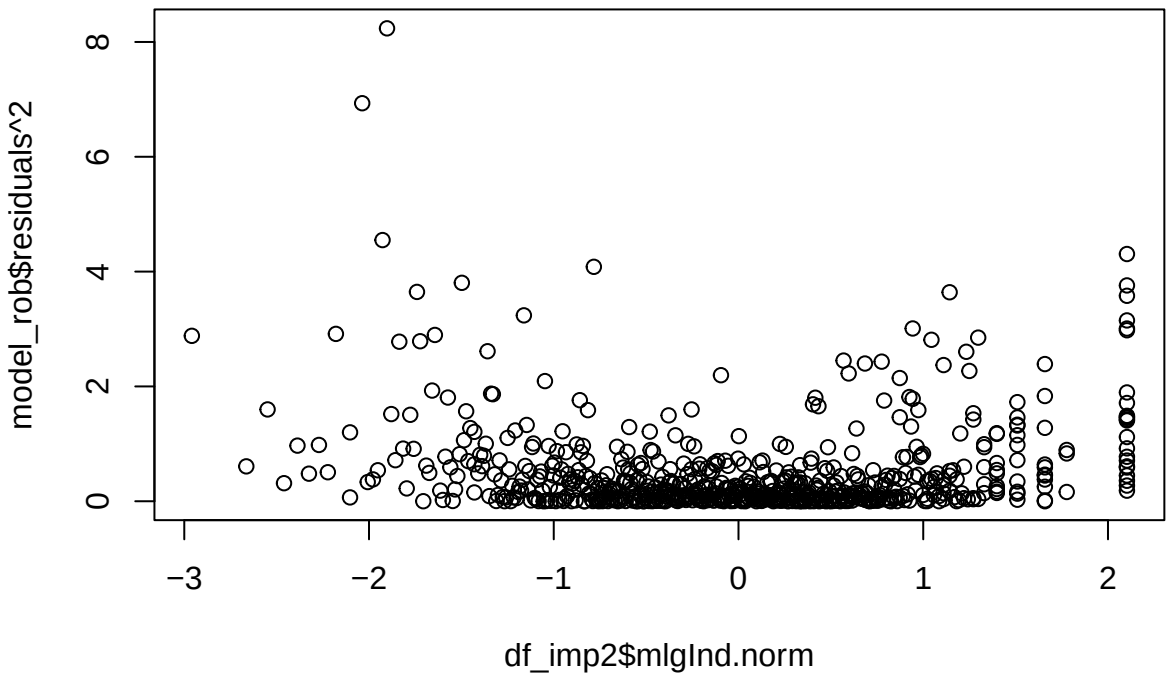
```
df_imp2 <- df_imp %>% filter(!(row.names(df_imp) %in% outlier_names[c(1,2,4)]))
df_imp2 <- weight_calc(df_imp2)

model_rob <- lm(mainFormula, weights = weight, data=df_imp2)

bptest(model_rob)
```

```
##
##  studentized Breusch-Pagan test
##
## data:  model_rob
## BP = 35.435, df = 26, p-value = 0.1026
```

```
plot(model_rob$residuals^2 ~ df_imp2$mlgInd.norm)
```




```
summary(model_rob)
```

```
##
## Call:
## lm(formula = mainFormula, data = df_imp2, weights = weight)
##
## Weighted Residuals:
##     Min      1Q  Median      3Q     Max
## -3.2761 -0.4543  0.0017  0.3984  2.4713
##
## Coefficients:
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)           -1.129526   0.211997  -5.328 1.39e-07 ***
## countryFrance          0.366667   0.136246   2.691 0.007312 **
## countryGermany         0.301595   0.139719   2.159 0.031267 *
## countryIreland         0.219759   0.137266   1.601 0.109892
## countryItaly           0.351227   0.139405   2.519 0.012004 *
## countryNetherlands     0.145513   0.135828   1.071 0.284451
## countryNorway          0.097510   0.138186   0.706 0.480675
## countrySpain           0.398818   0.137522   2.900 0.003864 **
## countrySweden          0.159515   0.138642   1.151 0.250362
## countrySwitzerland     0.075666   0.139959   0.541 0.588959
## countryUnited Kingdom  0.227068   0.137741   1.649 0.099756 .
## represent.bin1         0.331063   0.072164   4.588 5.43e-06 ***
## euInd.norm             0.137463   0.034736   3.957 8.46e-05 ***
```

```
## snaInd.norm            0.330105   0.029853  11.058  < 2e-16 ***
## effectiveIntOrg1       0.399319   0.067839   5.886 6.48e-09 ***
## effectiveDirRel1       0.369294   0.103137   3.581 0.000370 ***
## euFundImp.bin1         0.326697   0.086928   3.758 0.000187 ***
## polIdEu1              -0.011022   0.071956  -0.153 0.878308
## polIdCul1              0.321835   0.095003   3.388 0.000750 ***
## polIdIdeo1            -0.085357   0.073945  -1.154 0.248815
## LRCLeft                0.131327   0.066576   1.973 0.048988 *
## LRCRight              -0.143063   0.079611  -1.797 0.072818 .
## age                    0.002164   0.002457   0.881 0.378768
## genderMale            -0.056556   0.064201  -0.881 0.378705
## genderOther            0.117497   0.380126   0.309 0.757350
## eduLev.L               0.065829   0.071484   0.921 0.357465
## eduLev.Q              -0.000994   0.057043  -0.017 0.986103
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7188 on 618 degrees of freedom
## Multiple R-squared:  0.4854, Adjusted R-squared:  0.4638
## F-statistic: 22.42 on 26 and 618 DF,  p-value: < 2.2e-16
```

```
coeftest(model_rob, vcov=vcovHC(model_rob, type="HC2", cluster = ~country))
```

```
##
## t test of coefficients:
##
##                         Estimate Std. Error t value  Pr(>|t|)
## (Intercept)           -1.1295262  0.2207351 -5.1171 4.147e-07 ***
## countryFrance          0.3666675  0.1672609  2.1922 0.0287381 *
## countryGermany         0.3015946  0.1469435  2.0525 0.0405463 *
## countryIreland         0.2197595  0.1664650  1.3202 0.1872721
## countryItaly           0.3512269  0.1538494  2.2829 0.0227734 *
## countryNetherlands     0.1455129  0.1687390  0.8624 0.3888270
## countryNorway          0.0975101  0.1587443  0.6143 0.5392704
## countrySpain           0.3988183  0.2082340  1.9152 0.0559230 .
## countrySweden          0.1595150  0.1679194  0.9499 0.3425091
## countrySwitzerland     0.0756657  0.1329007  0.5693 0.5693324
## countryUnited Kingdom  0.2270675  0.1461317  1.5539 0.1207309
## represent.bin1         0.3310628  0.0848335  3.9025 0.0001057 ***
## euInd.norm             0.1374632  0.0413447  3.3248 0.0009372 ***
## snaInd.norm            0.3301047  0.0398944  8.2745 7.957e-16 ***
## effectiveIntOrg1       0.3993189  0.0778438  5.1297 3.889e-07 ***
## effectiveDirRel1       0.3692942  0.1148556  3.2153 0.0013711 **
## euFundImp.bin1         0.3266970  0.0914822  3.5712 0.0003830 ***
## polIdEu1              -0.0110221  0.0819241 -0.1345 0.8930196
## polIdCul1              0.3218352  0.1206111  2.6684 0.0078220 **
## polIdIdeo1            -0.0853567  0.0834841 -1.0224 0.3069769
## LRCLeft                0.1313267  0.0739450  1.7760 0.0762240 .
## LRCRight              -0.1430630  0.1006790 -1.4210 0.1558265
## age                    0.0021640  0.0025379  0.8527 0.3941672
## genderMale            -0.0565555  0.0749228 -0.7549 0.4506261
## genderOther            0.1174968  0.1456155  0.8069 0.4200359
## eduLev.L               0.0658294  0.0853391  0.7714 0.4407730
## eduLev.Q              -0.0009940  0.0665613 -0.0149 0.9880899
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

We repeat the same model with institution types:

```
model_rob_type <- lm(update(mainFormula, . ~ . -represent.bin + represent),
                     weights = weight, data=df_imp2)
summary(model_rob_type)
```

```
##
## Call:
## lm(formula = update(mainFormula, . ~ . - represent.bin + represent),
##     data = df_imp2, weights = weight)
##
## Weighted Residuals:
##     Min      1Q  Median      3Q     Max
## -3.2094 -0.4500 -0.0094  0.4095  2.4494
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)            -1.110774   0.212220  -5.234 2.28e-07 ***
## countryFrance           0.352531   0.136510   2.582 0.010040 *
## countryGermany          0.296964   0.139740   2.125 0.033974 *
## countryIreland          0.226879   0.138251   1.641 0.101294
## countryItaly            0.347671   0.139431   2.493 0.012911 *
## countryNetherlands      0.136022   0.135969   1.000 0.317517
## countryNorway           0.101621   0.138270   0.735 0.462655
## countrySpain            0.404572   0.137530   2.942 0.003387 **
## countrySweden           0.170962   0.138882   1.231 0.218795
## countrySwitzerland      0.071557   0.139938   0.511 0.609290
## countryUnited Kingdom   0.223128   0.137700   1.620 0.105660
## euInd.norm              0.133900   0.034831   3.844 0.000133 ***
## snaInd.norm             0.334413   0.029988  11.152  < 2e-16 ***
## effectiveIntOrg1        0.389282   0.068113   5.715 1.71e-08 ***
## effectiveDirRel1        0.382592   0.103451   3.698 0.000236 ***
## euFundImp.bin1          0.329584   0.087046   3.786 0.000168 ***
## polIdEu1               -0.021512   0.072399  -0.297 0.766463
## polIdCul1               0.319931   0.095195   3.361 0.000825 ***
## polIdIdeo1             -0.085383   0.073913  -1.155 0.248467
## LRCLeft                 0.135216   0.066699   2.027 0.043065 *
## LRCRight               -0.148064   0.079639  -1.859 0.063476 .
## age                     0.001647   0.002479   0.664 0.506666
## genderMale             -0.053007   0.064234  -0.825 0.409570
## genderOther             0.111140   0.380105   0.292 0.770085
## eduLev.L                0.065722   0.071624   0.918 0.359188
## eduLev.Q               -0.002417   0.057494  -0.042 0.966480
## representCOR            0.312096   0.111308   2.804 0.005208 **
## representCLRAE          0.556152   0.158494   3.509 0.000483 ***
## representOTHER          0.280704   0.092168   3.046 0.002422 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7185 on 616 degrees of freedom
```

```
## Multiple R-squared:  0.4876, Adjusted R-squared:  0.4643
## F-statistic: 20.93 on 28 and 616 DF,  p-value: < 2.2e-16
```

```
coeftest(model_rob_type, vcov=vcovHC(model_rob_type, type="HC2", cluster = ~country))
```

```
##
## t test of coefficients:
##
##                          Estimate Std. Error t value  Pr(>|t|)
## (Intercept)           -1.1107745  0.2204873 -5.0378 6.195e-07 ***
## countryFrance          0.3525308  0.1651352  2.1348 0.0331713 *
## countryGermany         0.2969642  0.1463552  2.0291 0.0428812 *
## countryIreland         0.2268792  0.1674204  1.3551 0.1758676
## countryItaly           0.3476706  0.1539323  2.2586 0.0242576 *
## countryNetherlands     0.1360217  0.1677946  0.8106 0.4178832
## countryNorway          0.1016207  0.1580944  0.6428 0.5206029
## countrySpain           0.4045722  0.2074608  1.9501 0.0516158 .
## countrySweden          0.1709622  0.1689623  1.0118 0.3120135
## countrySwitzerland     0.0715569  0.1320846  0.5418 0.5881861
## countryUnited Kingdom  0.2231285  0.1448767  1.5401 0.1240431
## euInd.norm             0.1338997  0.0414879  3.2274 0.0013153 **
## snaInd.norm            0.3344133  0.0399813  8.3642 4.050e-16 ***
## effectiveIntOrg1       0.3892822  0.0778810  4.9984 7.543e-07 ***
## effectiveDirRel1       0.3825916  0.1146535  3.3369 0.0008981 ***
## euFundImp.bin1         0.3295835  0.0923140  3.5702 0.0003844 ***
## polIdEu1              -0.0215124  0.0831577 -0.2587 0.7959573
## polIdCul1              0.3199306  0.1214549  2.6342 0.0086467 **
## polIdIdeo1            -0.0853828  0.0836534 -1.0207 0.3078101
## LRCLeft                0.1352165  0.0733336  1.8439 0.0656843 .
## LRCRight              -0.1480643  0.1006479 -1.4711 0.1417716
## age                    0.0016471  0.0025708  0.6407 0.5219495
## genderMale            -0.0530066  0.0749018 -0.7077 0.4794105
## genderOther            0.1111397  0.1502129  0.7399 0.4596540
## eduLev.L               0.0657223  0.0849411  0.7737 0.4393812
## eduLev.Q              -0.0024171  0.0680437 -0.0355 0.9716745
## representCOR           0.3120957  0.1468769  2.1249 0.0339942 *
## representCLRAE         0.5561525  0.1427545  3.8959 0.0001086 ***
## representOTHER         0.2807042  0.1140612  2.4610 0.0141283 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

We check if the difference between CoR an CLRA significant?:

```
df_imp2 <- df_imp2 %>% mutate(represent2 = relevel(represent, "CLRAE"))
model_rob_type2 <- lm(update(mainFormula, . ~ . -represent.bin + represent2),
                      weights = weight, data=df_imp2)
summary(model_rob_type2)
```

```
##
## Call:
## lm(formula = update(mainFormula, . ~ . - represent.bin + represent2),
##     data = df_imp2, weights = weight)
```

```
##
## Weighted Residuals:
##     Min      1Q  Median      3Q     Max
## -3.2094 -0.4500 -0.0094  0.4095  2.4494
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)            -0.554622   0.272981  -2.032 0.042610 *
## countryFrance           0.352531   0.136510   2.582 0.010040 *
## countryGermany          0.296964   0.139740   2.125 0.033974 *
## countryIreland          0.226879   0.138251   1.641 0.101294
## countryItaly            0.347671   0.139431   2.493 0.012911 *
## countryNetherlands      0.136022   0.135969   1.000 0.317517
## countryNorway           0.101621   0.138270   0.735 0.462655
## countrySpain            0.404572   0.137530   2.942 0.003387 **
## countrySweden           0.170962   0.138882   1.231 0.218795
## countrySwitzerland      0.071557   0.139938   0.511 0.609290
## countryUnited Kingdom   0.223128   0.137700   1.620 0.105660
## euInd.norm              0.133900   0.034831   3.844 0.000133 ***
## snaInd.norm             0.334413   0.029988  11.152  < 2e-16 ***
## effectiveIntOrg1        0.389282   0.068113   5.715 1.71e-08 ***
## effectiveDirRel1        0.382592   0.103451   3.698 0.000236 ***
## euFundImp.bin1          0.329584   0.087046   3.786 0.000168 ***
## polIdEu1               -0.021512   0.072399  -0.297 0.766463
## polIdCul1               0.319931   0.095195   3.361 0.000825 ***
## polIdIdeo1             -0.085383   0.073913  -1.155 0.248467
## LRCLeft                 0.135216   0.066699   2.027 0.043065 *
## LRCRight               -0.148064   0.079639  -1.859 0.063476 .
## age                     0.001647   0.002479   0.664 0.506666
## genderMale             -0.053007   0.064234  -0.825 0.409570
## genderOther             0.111140   0.380105   0.292 0.770085
## eduLev.L                0.065722   0.071624   0.918 0.359188
## eduLev.Q               -0.002417   0.057494  -0.042 0.966480
## represent2NONE         -0.556152   0.158494  -3.509 0.000483 ***
## represent2COR          -0.244057   0.182431  -1.338 0.181454
## represent2OTHER        -0.275448   0.174210  -1.581 0.114362
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7185 on 616 degrees of freedom
## Multiple R-squared:  0.4876, Adjusted R-squared:  0.4643
## F-statistic: 20.93 on 28 and 616 DF,  p-value: < 2.2e-16
```

```
coeftest(model_rob_type2, vcov=vcovHC(model_rob_type2, type="HC2", cluster = ~country))
```

```
##
## t test of coefficients:
##
##                          Estimate Std. Error t value  Pr(>|t|)
## (Intercept)            -0.5546220  0.2710372 -2.0463 0.0411514 *
## countryFrance           0.3525308  0.1651352  2.1348 0.0331713 *
## countryGermany          0.2969642  0.1463552  2.0291 0.0428812 *
## countryIreland          0.2268792  0.1674204  1.3551 0.1758676
## countryItaly            0.3476706  0.1539323  2.2586 0.0242576 *
```

```
## countryNetherlands     0.1360217  0.1677946  0.8106 0.4178832
## countryNorway          0.1016207  0.1580944  0.6428 0.5206029
## countrySpain           0.4045722  0.2074608  1.9501 0.0516158 .
## countrySweden          0.1709622  0.1689623  1.0118 0.3120135
## countrySwitzerland     0.0715569  0.1320846  0.5418 0.5881861
## countryUnited Kingdom  0.2231285  0.1448767  1.5401 0.1240431
## euInd.norm             0.1338997  0.0414879  3.2274 0.0013153 **
## snaInd.norm            0.3344133  0.0399813  8.3642 4.050e-16 ***
## effectiveIntOrg1       0.3892822  0.0778810  4.9984 7.543e-07 ***
## effectiveDirRel1       0.3825916  0.1146535  3.3369 0.0008981 ***
## euFundImp.bin1         0.3295835  0.0923140  3.5702 0.0003844 ***
## polIdEu1              -0.0215124  0.0831577 -0.2587 0.7959573
## polIdCul1              0.3199306  0.1214549  2.6342 0.0086467 **
## polIdIdeo1            -0.0853828  0.0836534 -1.0207 0.3078101
## LRCLeft                0.1352165  0.0733336  1.8439 0.0656843 .
## LRCRight              -0.1480643  0.1006479 -1.4711 0.1417716
## age                    0.0016471  0.0025708  0.6407 0.5219495
## genderMale            -0.0530066  0.0749018 -0.7077 0.4794105
## genderOther            0.1111397  0.1502129  0.7399 0.4596540
## eduLev.L               0.0657223  0.0849411  0.7737 0.4393812
## eduLev.Q              -0.0024171  0.0680437 -0.0355 0.9716745
## represent2NONE        -0.5561525  0.1427545 -3.8959 0.0001086 ***
## represent2COR         -0.2440567  0.1945058 -1.2548 0.2100443
## represent2OTHER       -0.2754483  0.1714458 -1.6066 0.1086502
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

In order to compare with different motivation levels, we create categories of roughly equal size, including three levels of motivation and actual experience.

```
df_imp3 <- df_imp2 %>%
  mutate(motiv_rep_com = case_when(
    represent.bin == 1 ~ "Experience",
    norepMotiv < 4 ~ "Low-mid motivation",
    norepMotiv == 4 ~ "High motivation",
    norepMotiv == 5 ~ "Very high motivation"
  ) %>% factor(levels = c(
    "Low-mid motivation", "High motivation", "Very high motivation", "Experience"
  ))) %>%
  drop_na(motiv_rep_com)
df_imp3 <- weight_calc(df_imp3)
```

We visualise distribution of unexplained variance across motivation levels and experience. For unexplained variation, we get residuals from the final model without the main predictor.

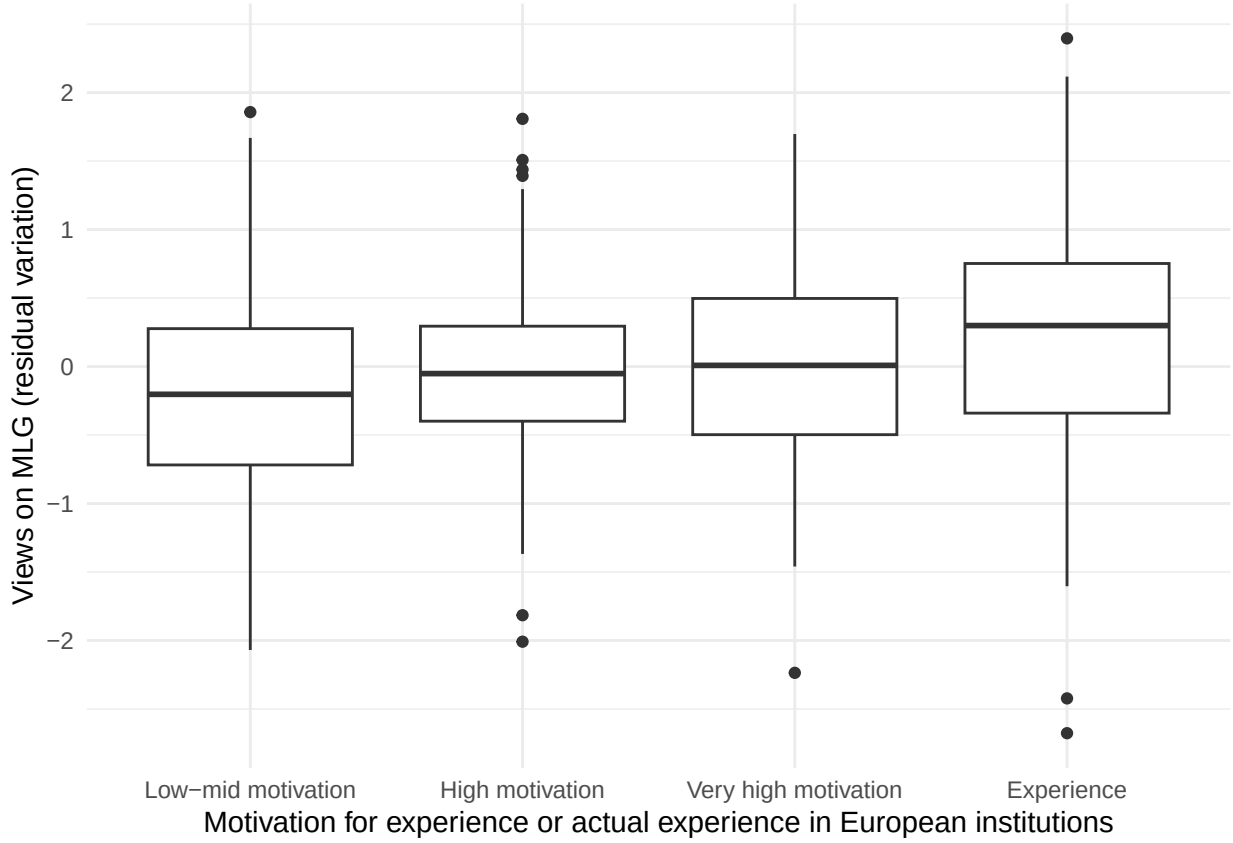
```
model_wo_iv <- lm(update(mainFormula, . ~ . -represent.bin),
                  weights=weight, data=df_imp3)
tbl_resids <- tibble(mlg = model_wo_iv$residuals, mot_rep = df_imp3$motiv_rep_com)

tbl_resids %>%
  group_by(mot_rep) %>%
  summarise(mean = mean(mlg))
```

```
## # A tibble: 4 x 2
```

```
##   mot_rep                 mean
##   <fct>                  <dbl>
## 1 Low-mid motivation   -0.164
## 2 High motivation      -0.0513
## 3 Very high motivation  0.0335
## 4 Experience            0.234
```

```
figure <- tbl_resids %>%
ggplot(aes(mot_rep, mlg)) +
  geom_boxplot() +
  theme_minimal() +
  labs(x = "Motivation for experience or actual experience in European institutions",
       y = "Views on MLG (residual variation)")
figure
```




```
ggsave(filename = "figure1_boxplot_exp.jpeg", plot = figure, device = "jpeg",
       width = 16, height = 12, units = "cm", dpi = 600)
```

We test the significance of the difference between very high motivation and experience

```
df_imp_himo <- df_imp2 %>%
  filter(represent.bin == 1 | norepMotiv > 4)
df_imp_himo <- weight_calc(df_imp_himo)
```

```r
model_comp <- lm(mainFormula, weights = weight, data = df_imp_himo)
summary(model_comp)
```

```
##
## Call:
## lm(formula = mainFormula, data = df_imp_himo, weights = weight)
##
## Weighted Residuals:
##      Min       1Q   Median       3Q      Max
## -3.10945 -0.52706 -0.01623  0.47770  1.96369
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)           -0.4010551  0.4065049  -0.987  0.32475
## countryFrance          0.6029742  0.2263656   2.664  0.00821 **
## countryGermany         0.4273161  0.2250092   1.899  0.05865 .
## countryIreland         0.1935253  0.2282243   0.848  0.39723
## countryItaly           0.3539807  0.2343874   1.510  0.13219
## countryNetherlands     0.2134362  0.2237853   0.954  0.34109
## countryNorway          0.2030281  0.2208729   0.919  0.35883
## countrySpain           0.5514879  0.2240447   2.462  0.01448 *
## countrySweden          0.1055014  0.2208624   0.478  0.63328
## countrySwitzerland     0.1494227  0.2237862   0.668  0.50491
## countryUnited Kingdom  0.2714052  0.2225209   1.220  0.22368
## represent.bin1         0.2229808  0.1008745   2.210  0.02794 *
## euInd.norm             0.1558559  0.0581736   2.679  0.00785 **
## snaInd.norm            0.2846953  0.0492049   5.786 2.05e-08 ***
## effectiveIntOrg1       0.2944451  0.1227289   2.399  0.01713 *
## effectiveDirRel1       0.1151243  0.2168064   0.531  0.59587
## euFundImp.bin1         0.2284231  0.1548823   1.475  0.14146
## polIdEu1              -0.0593533  0.1060055  -0.560  0.57602
## polIdCul1              0.4499251  0.1453142   3.096  0.00217 **
## polIdIdeo1            -0.1827705  0.1163896  -1.570  0.11754
## LRCLeft                0.0011339  0.1084214   0.010  0.99166
## LRCRight              -0.2680555  0.1328173  -2.018  0.04459 *
## age                    0.0005325  0.0040732   0.131  0.89610
## genderMale            -0.0888409  0.1103305  -0.805  0.42142
## eduLev.L               0.0576427  0.1083597   0.532  0.59521
## eduLev.Q               0.0204853  0.0906071   0.226  0.82131
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7694 on 262 degrees of freedom
## Multiple R-squared:  0.3633, Adjusted R-squared:  0.3026
## F-statistic: 5.981 on 25 and 262 DF,  p-value: 5.895e-15
```

```r
coeftest(model_comp, vcov=vcovHC(model_comp, type="HC2", cluster = ~country))
```

```
##
## t test of coefficients:
##
##                          Estimate  Std. Error t value  Pr(>|t|)
```

```
## (Intercept)           -0.40105514  0.40380472 -0.9932   0.32153
## countryFrance          0.60297423  0.29997990  2.0100   0.04545 *
## countryGermany         0.42731614  0.28283712  1.5108   0.13204
## countryIreland         0.19352535  0.30436380  0.6358   0.52544
## countryItaly           0.35398066  0.26593619  1.3311   0.18432
## countryNetherlands     0.21343618  0.32755112  0.6516   0.51522
## countryNorway          0.20302807  0.26983058  0.7524   0.45247
## countrySpain           0.55148788  0.30496464  1.8084   0.07170 .
## countrySweden          0.10550140  0.29005433  0.3637   0.71635
## countrySwitzerland     0.14942266  0.25439637  0.5874   0.55747
## countryUnited Kingdom  0.27140520  0.27499915  0.9869   0.32459
## represent.bin1         0.22298085  0.10808513  2.0630   0.04010 *
## euInd.norm             0.15585590  0.06236065  2.4993   0.01306 *
## snaInd.norm            0.28469534  0.05747985  4.9530 1.312e-06 ***
## effectiveIntOrg1       0.29444513  0.13998614  2.1034   0.03639 *
## effectiveDirRel1       0.11512431  0.20561342  0.5599   0.57602
## euFundImp.bin1         0.22842311  0.14125234  1.6171   0.10705
## polIdEu1              -0.05935329  0.11422673 -0.5196   0.60377
## polIdCul1              0.44992509  0.15591131  2.8858   0.00423 **
## polIdIdeo1            -0.18277048  0.10855448 -1.6837   0.09344 .
## LRCLeft                0.00113393  0.11499384  0.0099   0.99214
## LRCRight              -0.26805549  0.15147968 -1.7696   0.07796 .
## age                    0.00053245  0.00410690  0.1296   0.89694
## genderMale            -0.08884087  0.12877273 -0.6899   0.49086
## eduLev.L               0.05764268  0.12594667  0.4577   0.64757
## eduLev.Q               0.02048531  0.09281142  0.2207   0.82548
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```