Two-level MLM

Constantin Manuel Bosancianu[1]

WZB Berlin Social Science Center
Institutions and Political Inequality
January 18, 2021

The current tutorial goes through the steps of running in R a standard two-level mixed-effects model. The data set used throughout is the one you might be working on as you practice at home: the 6th wave of the European Social Survey, with a sample of 23 countries and a bit over 44,000 respondents.

First, read the data in `R` (since this is an `.Rnw` file, it will automatically use as working directory the folder in which it is placed). You can get a rough idea of what the variables in the data set measure, and how they are coded, by looking at the codebook for the ESS data, which is available in the `08-docs` sub-folder in the workshop folder I shared with you.

```
load("../02-data/ESS-practice.Rdata")

library(tidyverse)
ess <- ess %>%
  mutate(trustind = rowMeans(dplyr::select(.,ppltrst, pplfair, pplhlp),
                             na.rm = TRUE))
```

As dependent variable, I will be using a simple index of pro-social attitudes, which I've constructed here as a simple average of the three variables present in the data set: **ppltrst** ("most people can be trusted"), **pplfair** ("most people try to be fair"), and **pplhlp** ("most people try to be helpful"). More sophisticated measures can be used, of course, such as factor-based scales, but for our demonstration a simple average should be sufficient.

```
library(ggthemes)
ggplot(data=ess, aes(x=trustind)) +
  geom_histogram() +
  facet_wrap(~ cntry, ncol = 4) +
  theme_clean() +
  labs(x = "Pro-social attitudes",
```

[1] You can reach me at manuel.bosancianu@gmail.com. If you spot any mistakes I'd be grateful if you sent me an email pointing it out; I'll update the document and credit the help offered.

```
      y = "Frequency")
```

As can be quickly seen from Figure 1 on page 3, the distribution of the dependent variable is roughly normal in almost all countries in the sample. The only problem is the unexpectedly large number of answers of "5" on the scale for most countries. These are producing the spikes clearly seen in most panels of Figure 1. As there is not much that can be done to correct this problem, we are left with no choice but to proceed with the analysis.

1 Analyses

At the individual level, I will be interested in the relationship between education and pro-social attitudes. At the contextual level, I am keen on exploring how corruption impacts these same attitudes. To begin with, let's see whether there is any effect between education and pro-social attitudes, the shape of this effect, and whether it varies between national contexts.

```
ggplot(data = ess, aes(x = eduyrs, y=trustind)) +
  facet_wrap(~ cntry, ncol = 4) +
  stat_density_2d() +
  geom_smooth(method = "lm",
              se=FALSE,
              color = "red",
              size = 1.5) +
  theme_clean() +
  labs(x = "Years of education",
       y = "Pro-social attitudes")
```

The plot can be seen in Figure 2 on page 4 and presents us with considerable insights. To begin with, there is some variation in the relationship between education and pro-social attitudes. In some countries the relationship is clearly positive (BG, GER, SWE), while in others there appears to be no relationship (FIN, CZ, RUS). Finally, there is also the extreme case of Kosovo, where the relationship is negative.[2] More stories surface though. We can see from where the highest density of points is located that countries exhibit different average levels of pro-social attitudes. A clear distinction can be made between post-communist countries, at one end, and Scandinavian nations at the other end. The easiest spot in Figure 2 where this can be seen is the row that contains Kosovo, Norway and Poland, almost side by side. We see, then, that a post-communism dummy should likely be added to the analysis.

To begin with, let's run a simple null model—without any substantive predictors in the model. Before

[2] If this were an actual analysis, this is the stage where some deep thought would be put into whether Kosovo should be kept in the sample if it exhibits this unusual relationship. In my demo analysis I don't have to worry very much, as Kosovo will drop out of the sample due to missing information on the CPI and Gini index.

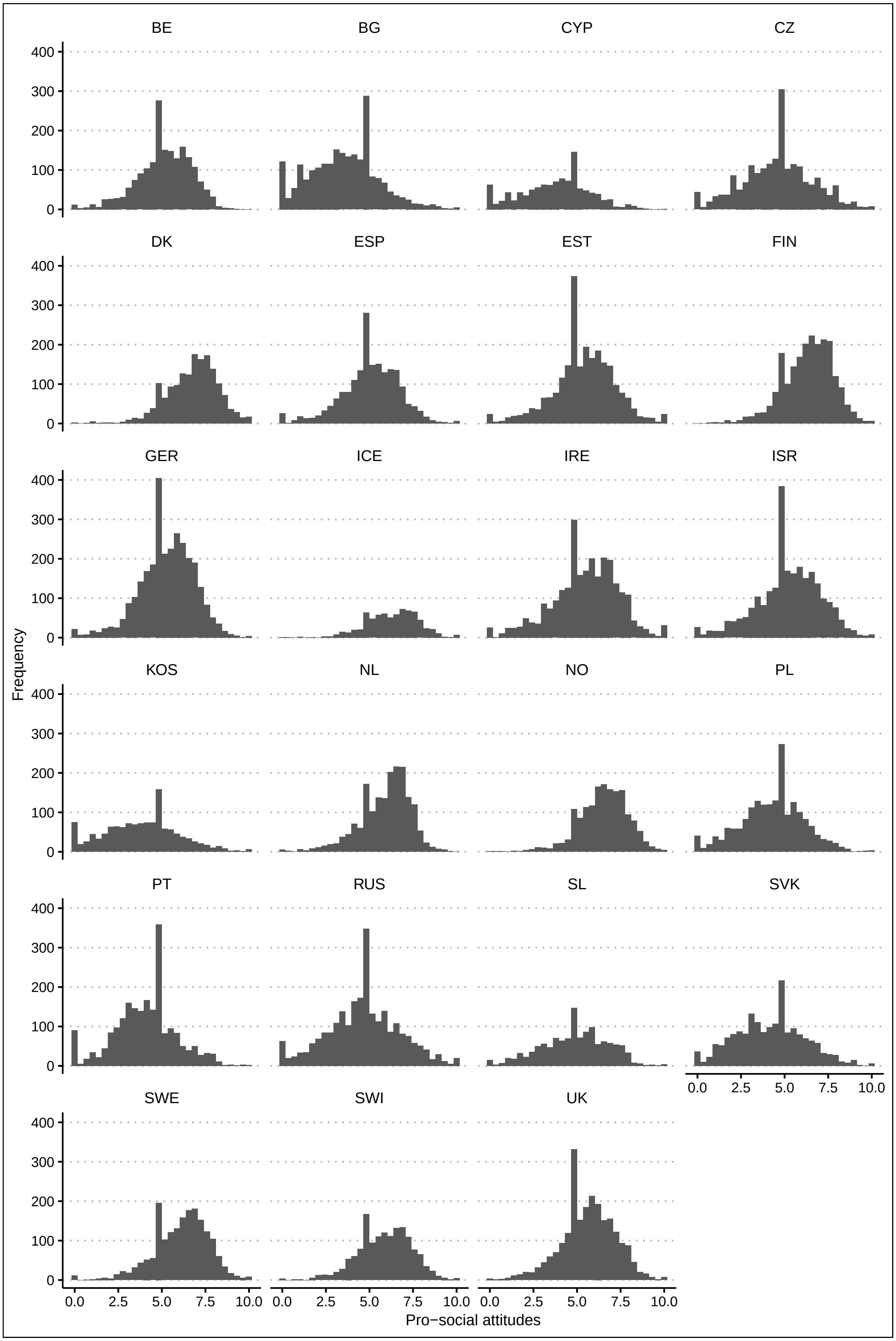

Figure 1: Distribution of pro-social attitudes per country

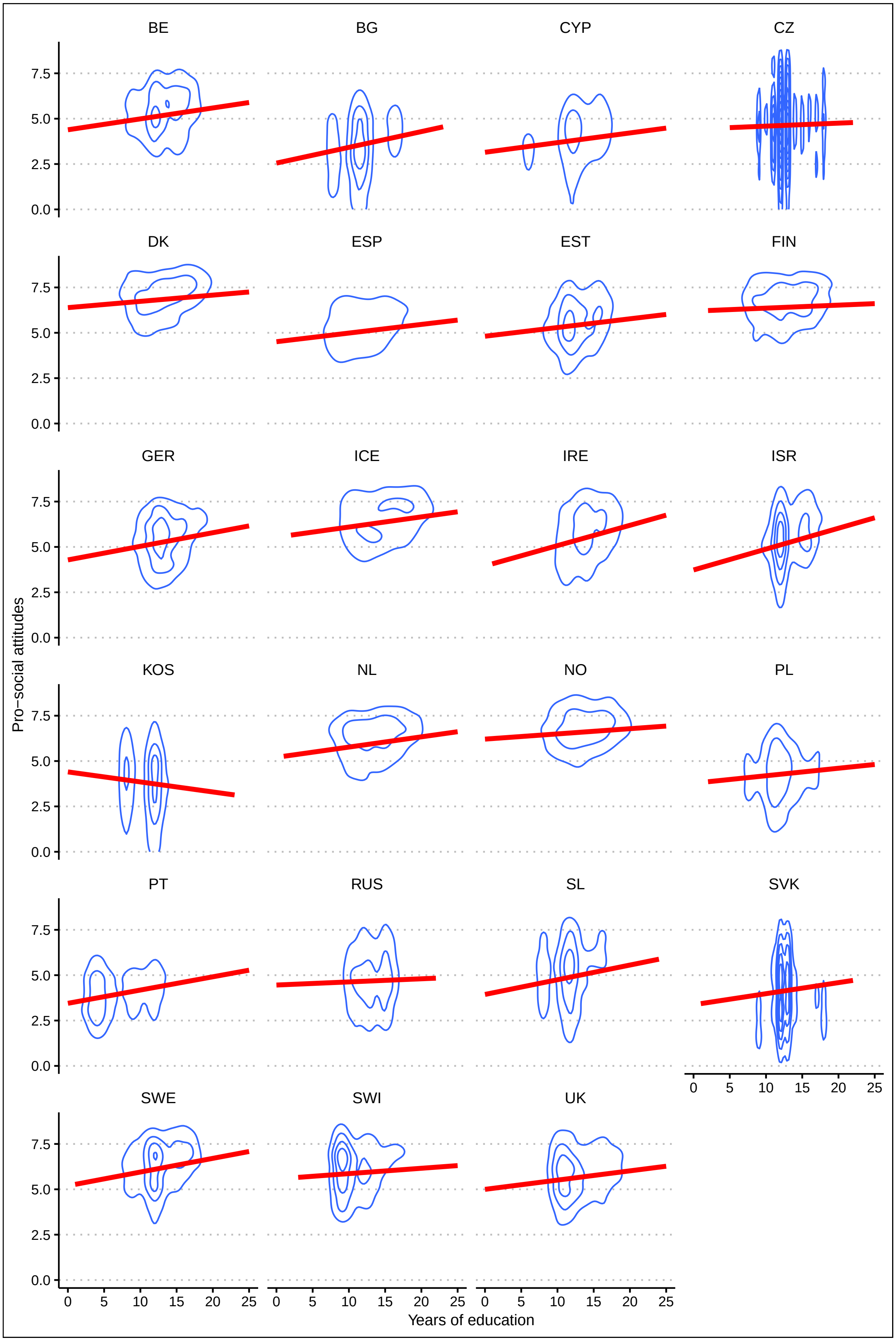

Figure 2: Relationship between education and pro-social attitudes

this, though, I have to center and standardize the variables I will be using in the analysis, and then clear all missing values.

```r
library(arm)

# Group-mean centering
ess <- ess %>%
  group_by(cntry) %>%
  mutate(age.cwc = arm::rescale(agea),
         gend.cwc = arm::rescale(male),
         edu.cwc = arm::rescale(eduyrs),
         inc.cwc = arm::rescale(hinctnta))

# Create country-level data
ess.agg <- ess %>%
  group_by(cntry) %>%
  distinct(gini_net, ti_cpi, postcom)
ess.agg <- data.frame(ess.agg)

ess.agg <- ess.agg %>%
  mutate(gini.cgm = arm::rescale(gini_net),
         postcom.cgm = arm::rescale(postcom),
         cpi.cgm = arm::rescale(ti_cpi)) %>%
  dplyr::select(-c(gini_net, ti_cpi, postcom))

ess <- left_join(ess, ess.agg, by = c("cntry"))
rm(ess.agg)

# Select only variables used and do listwise deletion
ess.sub <- ess %>%
  dplyr::select(cntry, age.cwc, gend.cwc, edu.cwc,
                inc.cwc, gini.cgm, postcom.cgm,
                cpi.cgm, trustind) %>%
  na.omit()

model0 <- lmer(trustind ~ 1 + (1 | cntry), # Specify model
               data = ess.sub, # Data to estimate model on
               na.action = na.omit, # Ignore missing cases (not needed here)
```

```r
                REML = TRUE) # REML estimation for parameters
summary(model0)
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: trustind ~ 1 + (1 | cntry)
   Data: ess.sub

REML criterion at convergence: 125382.3

Scaled residuals:
    Min      1Q  Median      3Q     Max 
-3.9242 -0.6198  0.0740  0.6961  3.6769 

Random effects:
 Groups   Name        Variance Std.Dev.
 cntry    (Intercept) 0.9326   0.9657  
 Residual             3.0644   1.7505  
Number of obs: 31648, groups:  cntry, 21

Fixed effects:
            Estimate Std. Error t value
(Intercept)    5.270      0.211   24.98
```

The model is estimated with REML, with individuals nested in countries. The ICC is the ratio of country intercept residual variance and total variance: $\frac{0.9326}{0.9326+3.0644}$= 0.233325. You can extract any needed quantities from the model, starting with fixed effects, random effects, SEs for fixed effects and random effects using the `fixef()`, `ranef()`, `se.fixef()` or `se.ranef()` functions from the **arm** package.

```r
fixef(model0)
se.fixef(model0)
```

```r
ranef(model0)
se.ranef(model0)
```

Keep in mind that these are not actual values of trust for each country; they are *deviations* of the countries from the overall level of trust. This is why some of them are negative, wile others are positive. You can also get a few model fit criteria with a few basic functions.

Table 1: First 10 random effects

Country	RE	SE
BE	-0.1101089	0.0431440
BG	-1.7068509	0.0397616
CYP	-1.4605981	0.0590348
CZ	-0.6583279	0.0475859
DK	1.5991362	0.0473416
ESP	-0.1883927	0.0452247
EST	0.1395813	0.0400206
FIN	1.1443538	0.0391797
GER	0.0471130	0.0351994
ICE	1.1233698	0.0713299

```
AIC(model0)
```

```
[1] 125388.3
```

```
BIC(model0)
```

```
[1] 125413.3
```

```
logLik(model0)
```

```
'log Lik.' -62691.13 (df=3)
```

My advice in terms of mixed-effects modeling is to start from simple models and gradually build up the model complexity. That way, if there are estimation warnings or errors, you can quickly diagnose which was the "guilty" predictor. If you start off from very complex specifications, it will be very difficult to figure out which of a number of fixed effects or random effects is the culprit. Here, I've started first with the individual-level predictors of pro-social attitudes.

```
model1 <- lmer(trustind ~ age.cwc + gend.cwc + edu.cwc +
                 inc.cwc + (1 | cntry),
               data = ess.sub,
               REML = TRUE)
summary(model1)
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: trustind ~ age.cwc + gend.cwc + edu.cwc + inc.cwc + (1 | cntry)
   Data: ess.sub

REML criterion at convergence: 124411.3
```

```
Scaled residuals:
    Min      1Q  Median      3Q     Max
-4.1698 -0.6249  0.0651  0.6674  3.9275

Random effects:
 Groups   Name        Variance Std.Dev.
 cntry    (Intercept) 0.9263   0.9625
 Residual             2.9699   1.7233
Number of obs: 31648, groups:  cntry, 21

Fixed effects:
            Estimate Std. Error t value
(Intercept)  5.26370    0.21028  25.032
age.cwc      0.29995    0.02092  14.335
gend.cwc    -0.13135    0.01961  -6.697
edu.cwc      0.38010    0.02148  17.698
inc.cwc      0.38925    0.02116  18.399

Correlation of Fixed Effects:
         (Intr) ag.cwc gnd.cw ed.cwc
age.cwc  -0.002
gend.cwc  0.000  0.017
edu.cwc  -0.002  0.242  0.030
inc.cwc   0.001  0.128 -0.100 -0.323
```

The results are consistent with what we would expect. A 2 SD increase in age would lead to a 0.3 increase in the level of pro-social attitudes, measured on a 0–10 scale. Men appear to be about -0.13 points less pro-social than women, on average. Both education and income have positive and statistically significant effects on the extent to which an individual exhibits pro-social attitudes. While the null model has an AIC of 1.2538825×10^5 and a logLikelihood of -6.269113×10^4, the AIC of the model with only individual-level predictors is 1.2442526×10^5, and the logLikelihood is -6.220563×10^4. With this difference between logLikelihoods there is little reason for a likelihood ratio test: the second model fits the data much better than the first one.[3]

We can now proceed to running a model with both individual-level and country-level predictors of pro-social attitudes.

[3] Remember, larger value for the logLikelihood denote a better fitting model. On the negative scale numbers that are closer to 0 are considered larger than numbers that are farther away from 0.

```r
model2 <- lmer(trustind ~ age.cwc + gend.cwc + edu.cwc +
                 inc.cwc + gini.cgm + postcom.cgm + cpi.cgm +
                 (1 | cntry),
               data = ess.sub,
               REML = TRUE)
summary(model2)
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: trustind ~ age.cwc + gend.cwc + edu.cwc + inc.cwc + gini.cgm +
    postcom.cgm + cpi.cgm + (1 | cntry)
   Data: ess.sub

REML criterion at convergence: 124384.2

Scaled residuals:
    Min      1Q  Median      3Q     Max
-4.1688 -0.6255  0.0648  0.6668  3.9278

Random effects:
 Groups   Name        Variance Std.Dev.
 cntry    (Intercept) 0.2531   0.5031
 Residual             2.9699   1.7233
Number of obs: 31648, groups:  cntry, 21

Fixed effects:
            Estimate Std. Error t value
(Intercept)  5.24718    0.11032  47.564
age.cwc      0.29991    0.02092  14.333
gend.cwc    -0.13138    0.01961  -6.699
edu.cwc      0.38006    0.02148  17.696
inc.cwc      0.38923    0.02116  18.398
gini.cgm     0.35080    0.29519   1.188
postcom.cgm  0.84480    0.43352   1.949
cpi.cgm      2.42905    0.47644   5.098

Correlation of Fixed Effects:
            (Intr) ag.cwc gnd.cw ed.cwc inc.cw gn.cgm pstcm.
age.cwc     -0.003
```

```
gend.cwc    -0.001  0.017
edu.cwc     -0.003  0.242  0.030
inc.cwc      0.001  0.128 -0.100 -0.323
gini.cgm    -0.010  0.000  0.000  0.000  0.000
postcom.cgm  0.014 -0.001  0.000 -0.002  0.000  0.408
cpi.cgm     -0.005  0.000  0.000 -0.003  0.001  0.616  0.832
```

Neither the effect of Gini or the post-communism dummy are statistically significant. However, in the case of post-communism, the effect is fairly strong which suggests that there might be something there. It's possible that this is a legacy of the Communist past, or something related to the cutthroat competitive environment immediately following the fall of Communism. On the other hand, the effect of corruption perceptions is clearly positive and statistically significant. A 2 SD increase in CPI (roughly speaking, 4.1 points on a 11 point scale, or about the difference between Romania and Norway in 2013) is associated with an average increase in pro-social attitudes of 2.43 points.[4] Does this new model fit the data better than the previous one, though?

```
anova(model1, model2)

Data: ess.sub
Models:
model1: trustind ~ age.cwc + gend.cwc + edu.cwc + inc.cwc + (1 | cntry)
model2: trustind ~ age.cwc + gend.cwc + edu.cwc + inc.cwc + gini.cgm +
model2:     postcom.cgm + cpi.cgm + (1 | cntry)
       npar    AIC    BIC logLik deviance  Chisq Df Pr(>Chisq)
model1    7 124400 124459 -62193   124386
model2   10 124376 124459 -62178   124356 30.509  3  1.079e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The likelihood ratio test can be used to compare the fit of the two models, to determine which is the better fitting model. There is a version of the test in the `lmtest` package, with the `lrtest()` command, but we suggest the standard `anova()` function. This has the benefit of correcting any user mistakes—if two models that differ in their fixed components are estimated with REML and then compared, `anova()` will re-estimate the models with FIML. With 10 estimated parameters (8 fixed effects and 2 random effects), Model 2 is shown to fit the data better than Model 1.

```
model3 <- lmer(trustind ~ age.cwc + gend.cwc + edu.cwc +
                 inc.cwc + gini.cgm + postcom.cgm + cpi.cgm +
                 (1 + edu.cwc | cntry),
```

[4] The CPI is reverse coded: increases in the score denote less corruption.

```
               data = ess.sub,
               REML = TRUE)
summary(model3)

Linear mixed model fit by REML ['lmerMod']
Formula: trustind ~ age.cwc + gend.cwc + edu.cwc + inc.cwc + gini.cgm +
    postcom.cgm + cpi.cgm + (1 + edu.cwc | cntry)
   Data: ess.sub

REML criterion at convergence: 124325.8

Scaled residuals:
    Min      1Q  Median      3Q     Max
-4.0568 -0.6222  0.0655  0.6676  3.9546

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 cntry    (Intercept) 0.2503   0.5003
          edu.cwc     0.0336   0.1833   -0.61
 Residual             2.9619   1.7210
Number of obs: 31648, groups:  cntry, 21

Fixed effects:
            Estimate Std. Error t value
(Intercept)  5.24721    0.10969  47.836
age.cwc      0.30639    0.02100  14.591
gend.cwc    -0.13358    0.01961  -6.812
edu.cwc      0.38774    0.04571   8.483
inc.cwc      0.38352    0.02117  18.119
gini.cgm     0.23670    0.24644   0.960
postcom.cgm  0.46086    0.36330   1.269
cpi.cgm      2.10215    0.39875   5.272

Correlation of Fixed Effects:
            (Intr) ag.cwc gnd.cw ed.cwc inc.cw gn.cgm pstcm.
age.cwc     -0.003
gend.cwc    -0.001  0.018
edu.cwc     -0.531  0.116  0.013
```

```
inc.cwc      0.001  0.125 -0.100 -0.152
gini.cgm    -0.009  0.001  0.004 -0.020 -0.004
postcom.cgm  0.011 -0.012  0.004 -0.014  0.002  0.403
cpi.cgm     -0.005 -0.010  0.004 -0.019  0.005  0.616  0.832
```

Does Model 3 fit the data better than Model 2? To rephrase it in a clearer way, do we have any reason to believe that the slope of education varies across countries?

```
anova(model2, model3)
```

```
Data: ess.sub
Models:
model2: trustind ~ age.cwc + gend.cwc + edu.cwc + inc.cwc + gini.cgm +
model2:     postcom.cgm + cpi.cgm + (1 | cntry)
model3: trustind ~ age.cwc + gend.cwc + edu.cwc + inc.cwc + gini.cgm +
model3:     postcom.cgm + cpi.cgm + (1 + edu.cwc | cntry)
       npar    AIC    BIC logLik deviance  Chisq Df Pr(>Chisq)
model2   10 124376 124459 -62178   124356
model3   12 124321 124422 -62149   124297 58.186  2  2.318e-13 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The results again suggest that the more complex model fits the data better. Adding two parameters to the estimation procedure (an extra random effect and the correlation between random effects) produced a significantly better fit to the data, as indicated by the likelihood ratio test we ran above. Finally, we can test to see whether corruption perceptions are a predictor of the relationship between education and pro-social attitudes

```
model4 <- lmer(trustind ~ age.cwc + gend.cwc + edu.cwc +
                 inc.cwc + gini.cgm + postcom.cgm + cpi.cgm +
                 edu.cwc*cpi.cgm + (1 + edu.cwc | cntry),
               data = ess.sub,
               REML = TRUE)
summary(model4)
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: trustind ~ age.cwc + gend.cwc + edu.cwc + inc.cwc + gini.cgm +
    postcom.cgm + cpi.cgm + edu.cwc * cpi.cgm + (1 + edu.cwc |      cntry)
   Data: ess.sub
```

```
REML criterion at convergence: 124328.8

Scaled residuals:
    Min      1Q  Median      3Q     Max 
-4.0567 -0.6216  0.0651  0.6674  3.9541 

Random effects:
 Groups   Name        Variance Std.Dev. Corr 
 cntry    (Intercept) 0.25400  0.5040        
          edu.cwc     0.03569  0.1889   -0.61
 Residual             2.96194  1.7210        
Number of obs: 31648, groups:  cntry, 21

Fixed effects:
                Estimate Std. Error t value
(Intercept)      5.24751    0.11050  47.488
age.cwc          0.30645    0.02100  14.592
gend.cwc        -0.13359    0.01961  -6.813
edu.cwc          0.38760    0.04681   8.280
inc.cwc          0.38351    0.02117  18.117
gini.cgm         0.23775    0.24653   0.964
postcom.cgm      0.46122    0.36342   1.269
cpi.cgm          2.07655    0.41688   4.981
edu.cwc:cpi.cgm  0.01937    0.09146   0.212

Correlation of Fixed Effects:
            (Intr) ag.cwc gnd.cw ed.cwc inc.cw gn.cgm pstcm. cp.cgm
age.cwc     -0.003                                                 
gend.cwc    -0.001  0.018                                          
edu.cwc     -0.540  0.114  0.013                                   
inc.cwc      0.001  0.124 -0.100 -0.149                            
gini.cgm    -0.008  0.001  0.004 -0.020 -0.004                     
postcom.cgm  0.011 -0.012  0.004 -0.014  0.002  0.403              
cpi.cgm     -0.008 -0.007  0.003 -0.010  0.001  0.586  0.795       
ed.cwc:cp.c  0.012 -0.007  0.001 -0.028  0.015  0.012  0.002 -0.291
```

Unfortunately, the model output indicates that corruption perceptions are not a moderator of the relationship between education and pro-social attitudes.[5] The coefficient for the interaction term is as far away from statistical significance as it can get. This is also indicated by the comparison of model fit statistics, which gives us little reason to increase the complexity of the model: the more intricate model does not fit the data significantly better than the less complex one.

```
anova(model3, model4)

Data: ess.sub
Models:
model3: trustind ~ age.cwc + gend.cwc + edu.cwc + inc.cwc + gini.cgm +
model3:     postcom.cgm + cpi.cgm + (1 + edu.cwc | cntry)
model4: trustind ~ age.cwc + gend.cwc + edu.cwc + inc.cwc + gini.cgm +
model4:     postcom.cgm + cpi.cgm + edu.cwc * cpi.cgm + (1 + edu.cwc |
model4:     cntry)
       npar    AIC    BIC logLik deviance Chisq Df Pr(>Chisq)
model3   12 124321 124422 -62149   124297
model4   13 124323 124432 -62149   124297 0.046  1     0.8302

# Clearly, model 4 does not fit the data better than
#   model 3.
```

2 Displaying model results

A standard way of presenting model results are regression comparison tables, easily generated through the `apsrtable`, `stargazer`, or `texreg` packages for R. In this example, I will use the latter package.

```
library(texreg)
# For LaTeX output
texreg(list(model0, model1, model2, model3, model4),
       digits=2,
       custom.model.names=c("Null model", "Model 1", "Model 2",
                            "Model 3", "Model 4"),
       custom.coef.names=c("(Intercept)", "Age (decades)",
                           "Male", "Education", "Income",
                           "Gini", "Post-communism",
                           "CPI index","Education*CPI"))
# For HTML output, that can be read into Word
htmlreg(list(model0, model1, model2, model3, model4),
```

[5] I increased the maximum number of iterations, as well as changed the optimizer to `bobyqa` ("bound optimization by quadratic approximation").

```r
        digits=2,
        custom.model.names=c("Null model", "Model 1", "Model 2",
                             "Model 3", "Model 4"),
        custom.coef.names=c("(Intercept)", "Age (decades)",
                            "Male", "Education", "Income",
                            "Gini", "Post-communism",
                            "CPI index","Education*CPI"))
# For output displayed in your R console
screenreg(list(model0, model1, model2, model3, model4),
          digits=2,
          custom.model.names=c("Null model", "Model 1", "Model 2",
                               "Model 3", "Model 4"),
          custom.coef.names=c("(Intercept)", "Age (decades)",
                              "Male", "Education", "Income",
                              "Gini", "Post-communism",
                              "CPI index","Education*CPI"))
```

As you can also see in the models presented in Table 2, you will always be required to report the fixed and random effects, as well as a few measures of model fit. Additionally, you have to report the sample sizes at all levels of the model. It is more a matter of preference if you also want to report the covariance between random effects, as I have done in the last row of the table.

Alternatively, you can present predicted values through the `effects` package, as seen in Figure 3.

```r
library(effects)
plot(Effect("cpi.cgm", model3),
  main="Predicted values plot",
  xlab="Corruption perceptions",
  ylab="Pro-social attitudes")
```

Had we obtained any significant results from the interaction model (in Model 4), we would probably have produced a marginal effects plot as well, to show how the effect of education varies depending on CPI.[6] Since we did not get a significant coefficient there is little to interpret, although for the sake of practice we can also produce this plot as well.

```r
quantile(ess.sub$cpi.cgm, c(0.1, 0.5, 0.9))
plot(effect("edu.cwc:cpi.cgm", model4,
```

[6]Remember that an interaction is symmetric, so the coefficient can be interpreted as how the effect of education varies across contexts with different CPIs, or how the effect of CPI is different for individuals with different levels of education. The interpretation depends mostly on how you have framed your theoretical questions.

Table 2: Statistical models

	Null model	Model 1	Model 2	Model 3	Model 4
(Intercept)	5.27***	5.26***	5.25***	5.25***	5.25***
	(0.21)	(0.21)	(0.11)	(0.11)	(0.11)
Age (decades)		0.30***	0.30***	0.31***	0.31***
		(0.02)	(0.02)	(0.02)	(0.02)
Male		−0.13***	−0.13***	−0.13***	−0.13***
		(0.02)	(0.02)	(0.02)	(0.02)
Education		0.38***	0.38***	0.39***	0.39***
		(0.02)	(0.02)	(0.05)	(0.05)
Income		0.39***	0.39***	0.38***	0.38***
		(0.02)	(0.02)	(0.02)	(0.02)
Gini			0.35	0.24	0.24
			(0.30)	(0.25)	(0.25)
Post-communism			0.84	0.46	0.46
			(0.43)	(0.36)	(0.36)
CPI index			2.43***	2.10***	2.08***
			(0.48)	(0.40)	(0.42)
Education*CPI					0.02
					(0.09)
AIC	125388.25	124425.26	124404.24	124349.85	124354.77
BIC	125413.34	124483.79	124487.87	124450.20	124463.48
Log Likelihood	−62691.13	−62205.63	−62192.12	−62162.92	−62164.39
Num. obs.	31648	31648	31648	31648	31648
Num. groups: cntry	21	21	21	21	21
Var: cntry (Intercept)	0.93	0.93	0.25	0.25	0.25
Var: Residual	3.06	2.97	2.97	2.96	2.96
Var: cntry edu.cwc				0.03	0.04
Cov: cntry (Intercept) edu.cwc				−0.06	−0.06

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

```
  xlevels=list(cpi.cgm=c(-0.85, 0.15, 0.65)),
  main="Marginal effects plot",
  xlab="Corruption perceptions",
  ylab="Effect of education on pro-social behavior")
```

As can be seen from Figure 4, there is no real distinction between the effect of education at different levels of CPI. In a sense this is what the non-significant coefficient of the interaction term told us as well.

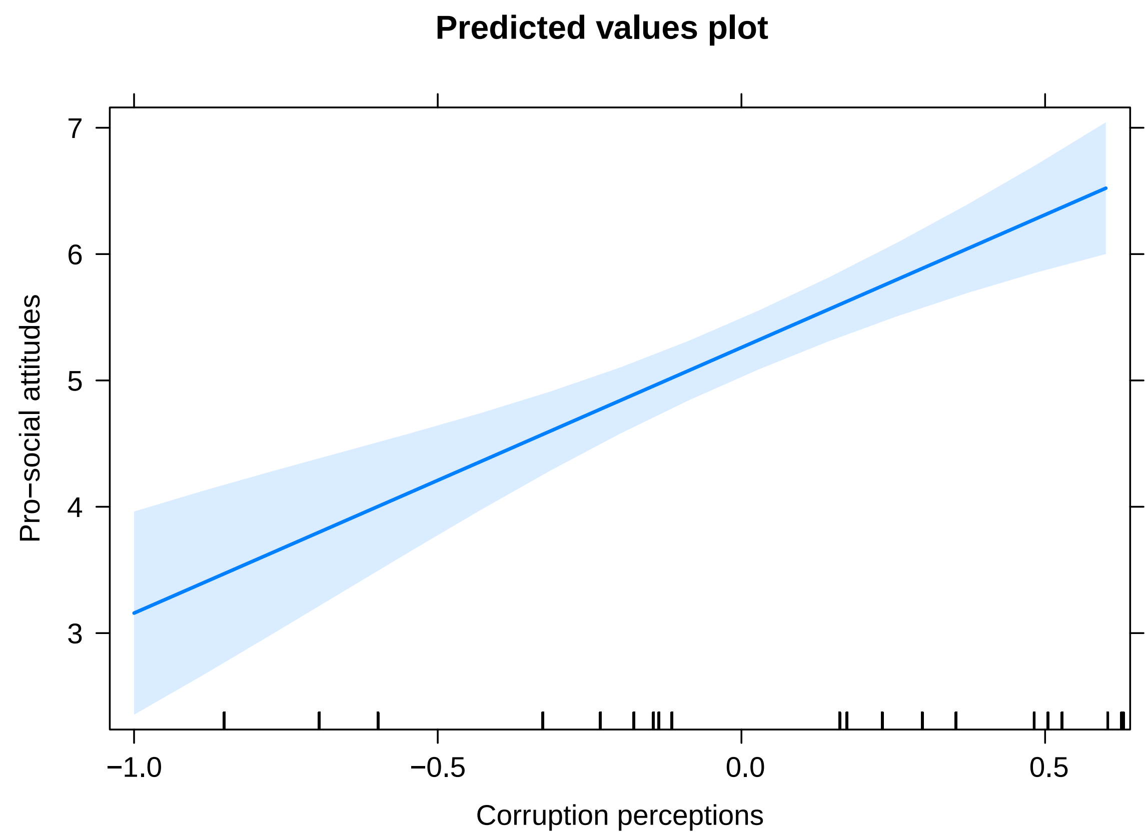

Figure 3: Predicted values with `effects` package

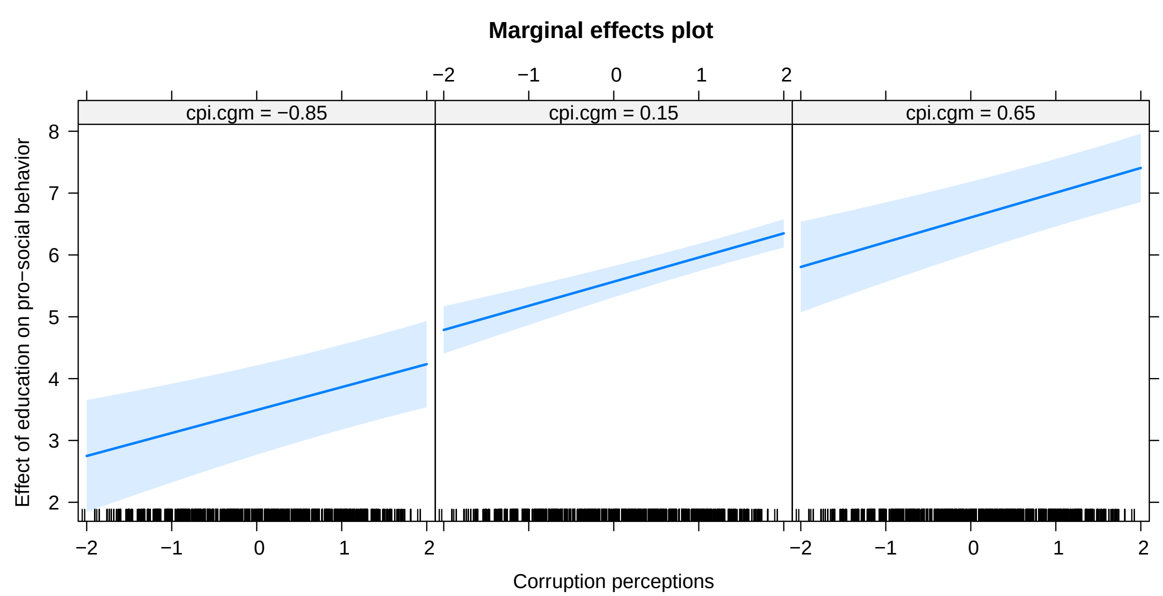

Figure 4: Marginal effects plot with `effects` package