

```
# Project1_507C_Kearns_GuggenmosEdits.R
# John Kearns (Edits by Greg Guggenmos)
# Date Created: 2022-11-19
# Last Updated: 2022-11-30
# Goal: Complete exercises recommended and extra analysis needed to inform Marlo G. Ball on
#       immigration attitudes in Switzerland

# set variables
master_dir = "C:/Users/16094/Documents/Year 1 - Princeton/Fall 2022/Quant Analysis/Project/"
do_folder = "C:/Users/16094/Documents/Year 1 - Princeton/Fall 2022/Quant Analysis/Project/Do/"
data_folder = "C:/Users/16094/Documents/Year 1 - Princeton/Fall 2022/Quant Analysis/Project/Data/"
results_folder = "C:/Users/16094/Documents/Year 1 - Princeton/Fall 2022/Quant Analysis/Project/Results/"
charts_folder = "C:/Users/16094/Documents/Year 1 - Princeton/Fall 2022/Quant Analysis/Project/Charts/"

# load packages
# install.packages("tidyverse")
# install.packages("estimatr")
# install.packages("pastecs")
# install.packages("texreg")
# install.packages("reactablefmtr")
# install.packages("plotrix")
#install.packages("logr")
#install.packages("reprex")

library(tidyverse)
library(estimatr)
library(pastecs)
#>
#> Attaching package: 'pastecs'
#> The following objects are masked from 'package:dplyr':
#>
#>     first, last
#> The following object is masked from 'package:tidyr':
#>
#>     extract
library(texreg)
```

```
#> Version: 1.38.6
#> Date: 2022-04-06
#> Author: Philip Leifeld (University of Essex)
#>
#> Consider submitting praise using the praise or praise_interactive functions.
#> Please cite the JSS article in your publications -- see citation("texreg").
#>
#> Attaching package: 'texreg'
#> The following object is masked from 'package:pastecs':
#>
#>     extract
#> The following object is masked from 'package:tidyverse':
#>
#>     extract
library(reactablefmtr)
#> Loading required package: reactable
#>
#> Attaching package: 'reactablefmtr'
#> The following object is masked from 'package:ggplot2':
#>
#>     margin
library(plotrix)
library(logr)
library(reprex)

# load data
data = haven::read_dta(paste0(data_folder,"swiss_passports.dta"))

# log using "swiss_passports.smcl", replace

# 1. Construct a table of summary statistics for the variables in the data set.
# summarize

# a. percent_novotes is the outcome variable of interest. What is its average value? What is its standard deviation?

summary(data$percent_novotes)
#>   Min. 1st Qu. Median Mean 3rd Qu. Max.
```

```

#>     4.00    29.00   42.00   40.35   52.00   88.00

data$percent_approved = ifelse(data$percent_novotes<50,1,0)
sd(data$percent_novotes)
#> [1] 14.78252

summary(data$percent_approved)
#>      Min. 1st Qu. Median      Mean 3rd Qu.      Max.
#> 0.000 0.000 1.000 0.685 1.000 1.000

#The average percentage of no votes is 40.4%. The standard deviation is 14.8%. The fraction of citizenship applicatio

# b. The various land_ variables are the explanatory variables of interest. What fraction of applications were from T
mean(data$land_ty==1|data$land_ceeu==1)
#> [1] 0.5612026
data = data %>%
  mutate(nationality=case_when(
    land_ty==1~"Turkey",
    land_ceeu==1~"Yugoslavia",
    land_asia==1~"Asia",
    land_seu==1~"Southern Europe",
    land_neu_poor==1~"NonEuropean Poor",
    land_other==1~"Other"
  ),
  decade=case_when(
    year80==1~"1980",
    year90==1~"1990",
    year00==1~"2000",
    year80==0&year90==0&year00==0~"1970"
  ))
)

print(data %>%
  group_by(nationality) %>%
  summarize(Share=n()/nrow(data)))
#> # A tibble: 6 x 2
#>   nationality     Share

```



```

#> shorter object length
#> `summarise()` has grouped output by 'decade'. You can override using the
#> `.`groups` argument.
#> # A tibble: 4 x 7
#>   decade   Asia `NonEuropean Poor` Other `Southern Europe` Turkey Yugoslavia
#>   <chr>    <dbl>          <dbl>    <dbl>      <dbl>        <dbl>
#> 1 1970    0.0155         0.00775  0.628     0.279  0.00775  0.0620
#> 2 1980    0.0176         0.0132   0.463     0.211  0.167    0.128
#> 3 1990    0.141       0.0298   0.0795    0.155  0.519    0.0755
#> 4 2000    0.0465         0.0204   0.0428    0.130  0.734    0.0260

# Overall, 49.7% are from Turkey, 6.3% from Yugoslavia, 16.6% from Southern Europe, 7.3% from Asia, 2.1% from non-Eur

# In the 70s, almost no immigrants from Turkey. Most were coming from Southern Europe and "other places". Turkish and

# There were 1,397 applications in this dataset with location data. 129 in 1970s, 227 in 80s, 503 in 90s, 538 in 2000

# Are there other noteworthy features of the data?

```

```

summary(data)
#> land_neu_poor      land_asia       year80       year90
#> Min. :0.00000  Min. :0.00000  Min. :0.0000  Min. :0.0000
#> 1st Qu.:0.00000  1st Qu.:0.00000  1st Qu.:0.0000  1st Qu.:0.0000
#> Median :0.00000  Median :0.00000  Median :0.0000  Median :0.0000
#> Mean   :0.02147  Mean   :0.07301  Mean   :0.1625  Mean   :0.3601
#> 3rd Qu.:0.00000  3rd Qu.:0.00000  3rd Qu.:0.0000  3rd Qu.:1.0000
#> Max.   :1.00000  Max.   :1.00000  Max.   :1.0000  Max.   :1.0000
#> 
#> year00      age40       age60       age60plus
#> Min. :0.0000  Min. :0.0000  Min. :0.0000  Min. :0.00000
#> 1st Qu.:0.0000 1st Qu.:0.0000  1st Qu.:0.0000  1st Qu.:0.00000
#> Median :0.0000  Median :0.0000  Median :0.0000  Median :0.00000
#> Mean   :0.3851  Mean   :0.4817  Mean   :0.3429  Mean   :0.02863
#> 3rd Qu.:1.0000  3rd Qu.:1.0000  3rd Qu.:1.0000  3rd Qu.:0.00000
#> Max.   :1.0000  Max.   :1.0000  Max.   :1.0000  Max.   :1.00000
#> 
#> male      kids       educ_middle   educ_high
#> Min. :0.0000  Min. :0.0000  Min. :0.0000  Min. :0.0000
#> 1st Qu.:0.0000 1st Qu.:0.0000  1st Qu.:0.0000  1st Qu.:0.0000

```

```
#> Median :1.0000  Median :1.0000  Median :1.0000  Median :0.0000
#> Mean    :0.7173  Mean    :0.5233  Mean    :0.5827  Mean    :0.1124
#> 3rd Qu.:1.0000 3rd Qu.:1.0000 3rd Qu.:1.0000 3rd Qu.:0.0000
#> Max.   :1.0000  Max.   :1.0000  Max.   :1.0000  Max.   :1.0000
#> married      lang_perfect     lang_good      lang_insufficient
#> Min.   :0.0000  Min.   :-1.000  Min.   :-1.000  Min.   :-1.0000
#> 1st Qu.:0.0000 1st Qu.:-1.000 1st Qu.:-1.000 1st Qu.:-1.0000
#> Median :1.0000  Median :-1.000  Median :-1.000  Median :-1.0000
#> Mean    :0.5648  Mean   :-0.267  Mean   :-0.5634 Mean   :-0.6034
#> 3rd Qu.:1.0000 3rd Qu.: 1.000 3rd Qu.: 0.0000 3rd Qu.: 0.0000
#> Max.   :1.0000  Max.   : 1.000  Max.   : 1.0000 Max.   : 1.0000
#> lang_NA       integration_NA  integration_integrated
#> Min.   :0.0000  Min.   :0.0000  Min.   :-1.0000
#> 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:-1.0000
#> Median :1.0000  Median :0.0000  Median : 0.0000
#> Mean    :0.6056  Mean   :0.3908  Mean   :-0.1818
#> 3rd Qu.:1.0000 3rd Qu.:1.0000 3rd Qu.: 0.0000
#> Max.   :1.0000  Max.   :1.0000  Max.   : 2.0000
#> integration_assimilated integration_adjusted integration_nodifference
#> Min.   :-1.0000  Min.   :-1.0000  Min.   :-1.000
#> 1st Qu.:-1.0000 1st Qu.:-1.0000 1st Qu.:-1.000
#> Median : 0.0000  Median : 0.0000  Median : 0.000
#> Mean    :-0.1038  Mean   :-0.3794  Mean   :-0.335
#> 3rd Qu.: 0.0000 3rd Qu.: 0.0000 3rd Qu.: 0.000
#> Max.   : 2.0000  Max.   : 1.0000  Max.   : 1.000
#> land_ty       land_ceeu      land_seu      land_other
#> Min.   :0.0000  Min.   :0.00000  Min.   :0.0000  Min.   :0.0000
#> 1st Qu.:0.0000 1st Qu.:0.00000 1st Qu.:0.0000 1st Qu.:0.0000
#> Median :0.0000  Median :0.00000  Median :0.0000  Median :0.0000
#> Mean    :0.4975  Mean   :0.06371  Mean   :0.1661  Mean   :0.1782
#> 3rd Qu.:1.0000 3rd Qu.:0.00000 3rd Qu.:0.0000 3rd Qu.:0.0000
#> Max.   :1.0000  Max.   :1.00000  Max.   :1.0000  Max.   :1.0000
#> skill_hi      skill_middle    born_ch      percent_novotes
#> Min.   :0.0000  Min.   :0.0000  Min.   :0.0000  Min.   : 4.00
#> 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:29.00
#> Median :0.0000  Median :0.0000  Median :0.0000  Median :42.00
#> Mean    :0.4216  Mean   :0.1654  Mean   :0.2076  Mean   :40.35
```

```

#> 3rd Qu.:1.0000 3rd Qu.:0.0000 3rd Qu.:0.0000 3rd Qu.:52.00
#> Max. :1.0000 Max. :1.0000 Max. :1.0000 Max. :88.00
#> percent_approved nationality decade
#> Min. :0.000 Length:1397 Length:1397
#> 1st Qu.:0.000 Class :character Class :character
#> Median :1.000 Mode :character Mode :character
#> Mean :0.685
#> 3rd Qu.:1.000
#> Max. :1.000

# Most immigrants are male. Lots of the language information is missing.

# 2. a. Construct table showing land_ and columns are number of obs, mean of percent_novotes, sd, standard error, and

print(data %>%
  group_by(nationality) %>%
  summarize(num_obs = n(),
           mean = mean(percent_novotes),
           st_dev = sd(percent_novotes),
           st_error = plotrix::std.error(percent_novotes),
           ci95_low = mean-1.96*st_error,
           ci95_high = mean+1.96*st_error))

#> # A tibble: 6 x 7
#>   nationality     num_obs   mean   st_dev   st_error   ci95_low   ci95_high
#>   <chr>         <int> <dbl> <dbl>    <dbl>    <dbl>    <dbl>
#> 1 Asia            102  39.1   10.3     1.02     37.1     41.1
#> 2 NonEuropean Poor      30  37.1   13.6     2.48     32.3     42.0
#> 3 Other            249  30.2   10.9     0.688    28.8     31.5
#> 4 Southern Europe      232  28.9   11.8     0.776    27.4     30.4
#> 5 Turkey            695  48.3   12.6     0.480    47.3     49.2
#> 6 Yugoslavia          89  39.4   13.6     1.45     36.6     42.3
# from this, we can see that Southern Europeans get the best votes while Turkey gets the worst.
# the Non-European poor have the highest variation in their acceptance.
# Southern Europe and Other are much more accepted than Asia, Turkey, Yugoslavia, and NonEuropean

mod1 = lm_robust(percent_novotes~land_ty+land_asia+land_ceeu+land_seu+land_neu_poor,data=data,se_type="stata")

```

```

# land_other is excluded so that there is no perfect collinearity

# from this we see how the significance of these differences from the mean varies by source country.
# Turkey has the largest magnitude, while Asia and Yugoslavia are around the same, relative to the "other" countries.
# Southern Europe get the best outcomes in votes, though the difference be SEU and "other" is not statistically signi

summary(mod1)
#>
#> Call:
#> lm_robust(formula = percent_novotes ~ land_ty + land_asia + land_ceeu +
#>     land_seu + land_neu_poor, data = data, se_type = "stata")
#>
#> Standard error type: HC1
#>
#> Coefficients:
#>
#>             Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper DF
#> (Intercept) 30.169    0.6879 43.855 1.598e-264   28.819 31.5181 1391
#> land_ty      18.086    0.8390 21.558 3.638e-89   16.440 19.7318 1391
#> land_asia     8.969    1.2297  7.293 5.055e-13    6.556 11.3808 1391
#> land_ceeu     9.258    1.5960  5.801 8.155e-09    6.127 12.3891 1391
#> land_seu     -1.268    1.0370 -1.223 2.217e-01   -3.302  0.7665 1391
#> land_neu_poor 6.965    2.5383  2.744 6.150e-03    1.985 11.9439 1391
#>
#> Multiple R-squared:  0.3284 ,   Adjusted R-squared:  0.326
#> F-statistic: 140.8 on 5 and 1391 DF,  p-value: < 2.2e-16

# 3. Add additional controls to regression
data = data %>%
  mutate(integration_integrated=ifelse(integration_integrated==1,NA,integration_integrated),
        integration_assimilated=ifelse(integration_assimilated==1,NA,integration_assimilated))

mod2 = lm_robust(percent_novotes~land_ty+land_asia+land_ceeu+land_seu+land_neu_poor+skill_hi+skill_middle,data=data,s
# b. Skill_ might differ across land_ variables because the educational systems are different across countries
# Also, Swiss people will understand the skills of countries more different than European countries.
# There is an argument for interaction terms,
# but at the very least high skilled people are more likely to be accepted because they can bring more to society.

```

```

summary(mod2)
#>
#> Call:
#> lm_robust(formula = percent_novotes ~ land_ty + land_asia + land_ceeu +
#>   land_seu + land_neu_poor + skill_hi + skill_middle, data = data,
#>   se_type = "stata")
#>
#> Standard error type: HC1
#>
#> Coefficients:
#>
#>             Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper DF
#> (Intercept) 32.234    0.9034 35.682 1.855e-198 30.462 34.0058 1389
#> land_ty      17.201    0.8971 19.174 6.992e-73   15.441 18.9607 1389
#> land_asia     8.108    1.2564  6.453 1.511e-10   5.643 10.5724 1389
#> land_ceeu     8.896    1.6076  5.534 3.747e-08   5.742 12.0492 1389
#> land_seu      -1.824   1.0632 -1.715 8.649e-02  -3.909  0.2618 1389
#> land_neu_poor 6.704    2.4920  2.690 7.225e-03   1.816 11.5927 1389
#> skill_hi      -2.105   0.7246 -2.904 3.739e-03  -3.526 -0.6831 1389
#> skill_middle   -3.347   1.0039 -3.334 8.779e-04  -5.317 -1.3779 1389
#>
#> Multiple R-squared:  0.3354 ,   Adjusted R-squared:  0.3321
#> F-statistic:  106 on 7 and 1389 DF,  p-value: < 2.2e-16

# Adding the skill_ variables increases the adjusted R2 by only a little bit
# However, the land_asia and land_ceeu coefficient change by a significant amount
# This would indicate that some of the acceptance these groups get is due to their skill levels

# c. male, kids, married, born_ch, age, and education all may be helpful
# male controls for some immigrant flows that may be older, more male, skewed towards family, have closer ties to Swi
mod3 = lm_robust(percent_novotes~land_ty+land_asia+land_ceeu+land_seu+land_neu_poor+year80+year90+year00+male+kids+ma

summary(mod3)
#>
#> Call:
#> lm_robust(formula = percent_novotes ~ land_ty + land_asia + land_ceeu +
#>   land_seu + land_neu_poor + year80 + year90 + year00 + male +
#>   kids + married + born_ch + age60 + age60plus + educ_high +

```

```

#>      skill_hi, data = data, se_type = "stata")
#>
#> Standard error type: HC1
#>
#> Coefficients:
#>
#>             Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper DF
#> (Intercept) 28.08649   1.4420 19.47801 7.588e-75 25.25782 30.9152 1380
#> land_ty     16.33408   1.1800 13.84278 6.652e-41 14.01935 18.6488 1380
#> land_asia    6.83350   1.4434  4.73422 2.425e-06  4.00195  9.6651 1380
#> land_ceeu    8.21849   1.6322  5.03525 5.402e-07  5.01665 11.4203 1380
#> land_seu    -1.46094   1.2332 -1.18470 2.363e-01 -3.88002  0.9581 1380
#> land_neu_poor 5.43620   2.7341  1.98830 4.698e-02  0.07277 10.7996 1380
#> year80      0.07498   1.3142  0.05706 9.545e-01 -2.50298  2.6529 1380
#> year90      3.46423   1.3898  2.49270 1.279e-02  0.73798  6.1905 1380
#> year00      2.04565   1.4519  1.40900 1.591e-01 -0.80242  4.8937 1380
#> male         1.36553   0.8306  1.64405 1.004e-01 -0.26383  2.9949 1380
#> kids         0.77784   1.2027  0.64674 5.179e-01 -1.58149  3.1372 1380
#> married      0.85134   1.1876  0.71684 4.736e-01 -1.47841  3.1811 1380
#> born_ch     -1.40307   0.9301 -1.50848 1.317e-01 -3.22768  0.4215 1380
#> age60        -0.22649   0.8294 -0.27307 7.848e-01 -1.85355  1.4006 1380
#> age60plus    -0.12391   1.5732 -0.07876 9.372e-01 -3.21013  2.9623 1380
#> educ_high   -1.98679   1.0832 -1.83420 6.684e-02 -4.11167  0.1381 1380
#> skill_hi    -0.18509   0.7431 -0.24908 8.033e-01 -1.64280  1.2726 1380
#>
#> Multiple R-squared:  0.3423 ,   Adjusted R-squared:  0.3347
#> F-statistic: 48.26 on 16 and 1380 DF,  p-value: < 2.2e-16

# coefficient on land_ty goes from 18.09 to 16.33
# coefficient on land_asia goes from 8.97 to 6.83
# coefficient on land_ceeu goes from 9.26 to 8.22
# coefficient on land_seu goes from -1.27 to -1.46
# coefficient on land_neu_poor goes from 6.97 to 5.44

# d. 1. How many observations on the language variable
sum(data$lang_NA==0)
#> [1] 551
# 551 out of 1397

```

```

sum(data$integration_NA==0)
#> [1] 851
# 851 out of 1397

# d. 2. Do the conclusions I reach in part c change if I restrict the samples?
mod4 = lm_robust(percent_novotes~land_ty+land_asia+land_ceeu+land_seu+land_neu_poor+year80+year90+year00+male+kids+ma
mod5 = lm_robust(percent_novotes~land_ty+land_asia+land_ceeu+land_seu+land_neu_poor+year80+year90+year00+male+kids+ma

summary(mod4)
#>
#> Call:
#> lm_robust(formula = percent_novotes ~ land_ty + land_asia + land_ceeu +
#>   land_seu + land_neu_poor + year80 + year90 + year00 + male +
#>   kids + married + born_ch + age60 + age60plus + educ_high +
#>   skill_hi, data = data %>% filter(lang_NA == 0), se_type = "stata")
#>
#> Standard error type: HC1
#>
#> Coefficients:
#>
#>             Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper DF
#> (Intercept) 27.45377 2.129 12.89794 2.516e-33 23.272 31.6351 534
#> land_ty      15.39028 2.060 7.47213 3.246e-13 11.344 19.4364 534
#> land_asia    6.83965 2.398 2.85278 4.502e-03 2.130 11.5494 534
#> land_ceeu    2.88746 3.033 0.95200 3.415e-01 -3.071 8.8456 534
#> land_seu     -2.49620 2.024 -1.23317 2.181e-01 -6.473 1.4802 534
#> land_neu_poor 4.06837 4.366 0.93190 3.518e-01 -4.508 12.6444 534
#> year80       0.83405 2.167 0.38494 7.004e-01 -3.422 5.0903 534
#> year90       8.84449 2.053 4.30858 1.956e-05 4.812 12.8770 534
#> year00      10.00967 2.145 4.66656 3.877e-06 5.796 14.2233 534
#> male          -0.06454 1.019 -0.06336 9.495e-01 -2.065 1.9363 534
#> kids          -1.83614 1.814 -1.01246 3.118e-01 -5.399 1.7264 534
#> married       1.70543 1.667 1.02313 3.067e-01 -1.569 4.9799 534
#> born_ch      -2.69213 1.177 -2.28702 2.259e-02 -5.005 -0.3797 534
#> age60         -1.13350 1.172 -0.96733 3.338e-01 -3.435 1.1684 534
#> age60plus     -2.93614 2.751 -1.06730 2.863e-01 -8.340 2.4679 534
#> educ_high     1.41142 1.680 0.84003 4.013e-01 -1.889 4.7120 534

```

```

#> skill_hi      -0.23639      1.084 -0.21806 8.275e-01   -2.366   1.8931 534
#>
#> Multiple R-squared:  0.487 , Adjusted R-squared:  0.4717
#> F-statistic: 37.47 on 16 and 534 DF,  p-value: < 2.2e-16
# the magnitudes are lower for those where language is available, but especially for those from Yugoslavia

summary(mod5)
#>
#> Call:
#> lm_robust(formula = percent_novotes ~ land_ty + land_asia + land_ceeu +
#>     land_seu + land_neu_poor + year80 + year90 + year00 + male +
#>     kids + married + born_ch + age60 + age60plus + educ_high +
#>     skill_hi, data = data %>% filter(integration_NA == 0), se_type = "stata")
#>
#> Standard error type: HC1
#>
#> Coefficients:
#>
#>             Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper DF
#> (Intercept) 28.33369  1.4251 19.8814 2.689e-72  25.5364 31.1310 834
#> land_ty      16.07431  1.2909 12.4523 9.106e-33  13.5406 18.6080 834
#> land_asia    6.80977  1.5573  4.3728 1.382e-05  3.7531  9.8665 834
#> land_ceeu    7.87940  1.8601  4.2360 2.529e-05  4.2284 11.5304 834
#> land_seu     -0.73781  1.3414 -0.5500 5.825e-01 -3.3708  1.8951 834
#> land_neu_poor 10.66587  1.8327  5.8198 8.398e-09  7.0687 14.2631 834
#> year80       1.28708  1.3531  0.9512 3.418e-01 -1.3688  3.9430 834
#> year90       5.80397  1.3702  4.2359 2.530e-05  3.1145  8.4934 834
#> year00       5.84409  1.4448  4.0449 5.720e-05  3.0082  8.6799 834
#> male          1.27753  0.8180  1.5617 1.187e-01 -0.3281  2.8832 834
#> kids          -1.11586  1.1600 -0.9620 3.363e-01 -3.3927  1.1610 834
#> married       1.91662  1.1469  1.6712 9.506e-02 -0.3345  4.1677 834
#> born_ch      -2.67022  0.9166 -2.9132 3.673e-03 -4.4693 -0.8711 834
#> age60         -0.30580  0.8376 -0.3651 7.152e-01 -1.9499  1.3383 834
#> age60plus    -1.58199  1.7776 -0.8899 3.738e-01 -5.0711  1.9072 834
#> educ_high    -0.84604  1.1017 -0.7679 4.428e-01 -3.0085  1.3165 834
#> skill_hi      0.02408  0.7546  0.0319 9.746e-01 -1.4571  1.5053 834
#>
#> Multiple R-squared:  0.4872 ,   Adjusted R-squared:  0.4774

```

```

#> F-statistic: 51.74 on 16 and 834 DF,  p-value: < 2.2e-16
# in this subset, the neu_poor coefficient doubles in size, but the rest are essentially unchanged

# d.3. add language variables as additional controls
data = data %>%
  mutate_at(vars(lang_perfect:lang_insufficient), funs(ifelse(.===-1, NA, .)))
#> Warning: `funs()` was deprecated in dplyr 0.8.0.
#> Please use a list of either functions or lambdas:
#>
#> # Simple named list:
#> list(mean = mean, median = median)
#>
#> # Auto named with `tibble::lst()`:
#> tibble::lst(mean, median)
#>
#> # Using lambdas
#> list(~ mean(., trim = .2), ~ median(., na.rm = TRUE))
#> This warning is displayed once every 8 hours.
#> Call `lifecycle::last_lifecycle_warnings()` to see where this warning was generated.

mod6 = lm_robust(percent_novotes~land_ty+land_asia+land_ceeu+land_seu+land_neu_poor+year80+year90+year00+male+kids+ma
summary(mod6)
#>
#> Call:
#> lm_robust(formula = percent_novotes ~ land_ty + land_asia + land_ceeu +
#>   land_seu + land_neu_poor + year80 + year90 + year00 + male +
#>   kids + married + born_ch + age60 + age60plus + educ_high +
#>   skill_hi + lang_good + lang_perfect, data = data %>% filter(integration_NA ==
#>   0), se_type = "stata")
#>
#> Standard error type: HC1
#>
#> Coefficients:
#>             Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper DF
#> (Intercept)  34.0215    3.133 10.8607 1.622e-24   27.865  40.1781 442
#> land_ty      16.4886    2.044  8.0680 6.799e-15   12.472  20.5052 442
#> land_asia     6.1449    2.445  2.5132 1.232e-02    1.340  10.9502 442

```

```

#> land_ceeu      2.7255    3.047  0.8944 3.716e-01   -3.263   8.7142 442
#> land_seu      -1.4490   1.979 -0.7321 4.645e-01   -5.339   2.4409 442
#> land_neu_poor 12.5806   2.796  4.4993 8.726e-06   7.085   18.0759 442
#> year80         1.6200    2.038  0.7951 4.270e-01   -2.384   5.6244 442
#> year90         8.9009    1.995  4.4621 1.031e-05   4.980   12.8213 442
#> year00         8.4784    2.087  4.0625 5.744e-05   4.377   12.5800 442
#> male            0.4468    1.052  0.4248 6.712e-01   -1.620   2.5139 442
#> kids             -3.2080   1.729 -1.8553 6.422e-02   -6.606   0.1903 442
#> married          2.1600    1.638  1.3189 1.879e-01   -1.059   5.3787 442
#> born_ch          -2.9139   1.179 -2.4722 1.380e-02   -5.230   -0.5974 442
#> age60            -1.3807   1.137 -1.2147 2.251e-01   -3.615   0.8532 442
#> age60plus        -5.4934   2.020 -2.7195 6.796e-03   -9.463   -1.5233 442
#> educ_high         1.6953    1.727  0.9815 3.269e-01   -1.699   5.0898 442
#> skill_hi          -0.1330   1.095 -0.1215 9.033e-01   -2.284   2.0183 442
#> lang_good          -5.2381   2.252 -2.3264 2.045e-02   -9.663   -0.8129 442
#> lang_perfect       -6.6504   2.169 -3.0657 2.305e-03   -10.914  -2.3870 442
#>
#> Multiple R-squared:  0.5453 ,     Adjusted R-squared:  0.5268
#> F-statistic: 36.38 on 18 and 442 DF,  p-value: < 2.2e-16

```

```

# the poor outcomes for NEU Poor immigrants becomes even more clear. The coefficient on the others except Yugoslavia
# R2 is now 0.53
# the language variables are significant and are consistent with my priors.
# Immigrants are more accepted if they speak the language better.

```

```

# d.4. Add the integration variables
data = data %>%
  mutate_at(vars(integration_integrated:integration_nodifference), funs(ifelse(.== -1, NA, .)))

View(filter(data, integration_assimilated == 1.5))
mod7 = lm_robust(percent_novotes~land_ty+land_asia+
                  land_ceeu+land_seu+land_neu_poor+year80+
                  year90+year00+male+kids+married+born_ch+
                  age60+age60plus+educ_high+skill_hi+
                  lang_good+lang_perfect+factor(integration_integrated)+
                  factor(integration_assimilated)+integration_adjusted+
                  integration_nodifference, data = data %>% filter(integration_NA==0), se_type="stata")

```

```

summary(mod7)
#>
#> Call:
#> lm_robust(formula = percent_novotes ~ land_ty + land_asia + land_ceeu +
#>   land_seu + land_neu_poor + year80 + year90 + year00 + male +
#>   kids + married + born_ch + age60 + age60plus + educ_high +
#>   skill_hi + lang_good + lang_perfect + factor(integration_integrated) +
#>   factor(integration_assimilated) + integration_adjusted +
#>
#> Standard error type: HC1
#>
#> Coefficients:
#>                               Estimate Std. Error t value Pr(>|t|)
#> (Intercept)                 33.7152   3.0446 11.0738 2.813e-25
#> land_ty                      17.3411   1.9568  8.8621 2.028e-17
#> land_asia                     6.6341   2.5374  2.6145 9.245e-03
#> land_ceeu                     5.7281   2.5693  2.2295 2.629e-02
#> land_seu                     -1.4611   1.9100 -0.7650 4.447e-01
#> land_neu_poor                  13.5970   2.4777  5.4878 6.924e-08
#> year80                        0.8469   2.1237  0.3988 6.902e-01
#> year90                        9.8630   2.1292  4.6322 4.785e-06
#> year00                        9.4053   2.2002  4.2748 2.353e-05
#> male                           1.1318   1.0001  1.1316 2.584e-01
#> kids                            -2.0884   1.3217 -1.5801 1.148e-01
#> married                         1.5869   1.3569  1.1695 2.428e-01
#> born_ch                          -3.5065   1.0674 -3.2850 1.102e-03
#> age60                           -1.3539   0.9976 -1.3572 1.754e-01
#> age60plus                       -4.6952   2.3806 -1.9723 4.921e-02
#> educ_high                        2.0588   1.3544  1.5201 1.292e-01
#> skill_hi                          0.1668   1.0047  0.1660 8.682e-01
#> lang_good                         -3.9988   2.2045 -1.8139 7.038e-02
#> lang_perfect                      -5.2630   2.1420 -2.4571 1.440e-02
#> factor(integration_integrated)1    -5.1211   1.0493 -4.8807 1.487e-06
#> factor(integration_integrated)2    -1.2908   1.6630 -0.7762 4.381e-01
#> factor(integration_assimilated)1    -2.8079   1.6637 -1.6878 9.217e-02
#> factor(integration_assimilated)1.5 -24.1165   2.4635 -9.7894 1.390e-20
#> factor(integration_assimilated)2      0.5460   1.7126  0.3188 7.500e-01

```

```

#> integration_adjusted           -4.8069    2.0649 -2.3279 2.037e-02
#> integration_nodifference      -3.4525    1.3442 -2.5685 1.055e-02
#>
#> (Intercept)                   27.7312   39.69909 435
#> land_ty                        13.4952   21.18699 435
#> land_asia                      1.6471   11.62114 435
#> land_ceeu                      0.6784   10.77781 435
#> land_seu                       -5.2152   2.29289 435
#> land_neu_poor                  8.7273   18.46678 435
#> year80                         -3.3271   5.02101 435
#> year90                         5.6782   14.04778 435
#> year00                         5.0810   13.72956 435
#> male                            -0.8339   3.09739 435
#> kids                            -4.6860   0.50928 435
#> married                         -1.0800   4.25379 435
#> born_ch                         -5.6045   -1.40854 435
#> age60                           -3.3145   0.60675 435
#> age60plus                      -9.3741   -0.01638 435
#> educ_high                       -0.6032   4.72077 435
#> skill_hi                        -1.8079   2.14150 435
#> lang_good                        -8.3317   0.33409 435
#> lang_perfect                     -9.4729   -1.05307 435
#> factor(integration_integrated)1 -7.1833   -3.05882 435
#> factor(integration_integrated)2 -4.5594   1.97777 435
#> factor(integration_assimilated)1 -6.0777   0.46191 435
#> factor(integration_assimilated)1.5 -28.9585  -19.27462 435
#> factor(integration_assimilated)2 -2.8200   3.91189 435
#> integration_adjusted           -8.8653   -0.74855 435
#> integration_nodifference      -6.0944   -0.81060 435
#>
#> Multiple R-squared:  0.6528 ,   Adjusted R-squared:  0.6329
#> F-statistic: 44 on 25 and 435 DF,  p-value: < 2.2e-16
# I think integration assimilated might have a few errors, will be worth cleaning up

# Coefficients for turkey and NEU get larger.
# Coefficients on the integration variables are almost all significant with negative coefficients
# This fits with my intuition that as integration and assimilation improve, people will be more willing to accept the

```

```

# 4. Do the land_ coefficients depend on occupation skills or educational level
mod8 = lm_robust(percent_novotes~land_ty+land_asia+
                  land_ceeu+land_seu+land_neu_poor+
                  educ_high+educ_middle+skill_middle+skill_hi,data=data %>% filter(integration_NA==0),se_type="stata"

summary(mod8)
#>
#> Call:
#> lm_robust(formula = percent_novotes ~ land_ty + land_asia + land_ceeu +
#>     land_seu + land_neu_poor + educ_high + educ_middle + skill_middle +
#>     skill_hi, data = data %>% filter(integration_NA == 0), se_type = "stata")
#>
#> Standard error type:  HC1
#>
#> Coefficients:
#>             Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper DF
#> (Intercept) 32.1633   1.1204 28.7063 7.215e-127  29.964 34.36250 841
#> land_ty      19.2622   0.9866 19.5243 2.662e-70   17.326 21.19865 841
#> land_asia    10.1351   1.3881  7.3013 6.610e-13   7.410 12.85966 841
#> land_ceeu    10.3641   1.8610  5.5691 3.445e-08   6.711 14.01678 841
#> land_seu     -0.5063   1.1599 -0.4365 6.626e-01  -2.783 1.77034 841
#> land_neu_poor 13.5213   2.1665  6.2410 6.886e-10   9.269 17.77370 841
#> educ_high     1.2054   1.3147  0.9169 3.595e-01  -1.375 3.78590 841
#> educ_middle    0.1368   0.7487  0.1827 8.551e-01  -1.333 1.60625 841
#> skill_middle   -3.0876   1.1897 -2.5953 9.614e-03  -5.423 -0.75253 841
#> skill_hi       -1.4916   0.7469 -1.9972 4.613e-02  -2.958 -0.02567 841
#>
#> Multiple R-squared:  0.4663 ,   Adjusted R-squared:  0.4606
#> F-statistic: 90.23 on 9 and 841 DF,  p-value: < 2.2e-16
# it does appear that these variables are important

mod9 = lm_robust(percent_novotes~land_ty*educ_high+land_ty*educ_middle+land_ty*skill_middle+land_ty*skill_hi+
                  land_asia*educ_high+land_asia*educ_middle+land_asia*skill_middle+land_asia*skill_hi+
                  land_ceeu*educ_high+land_ceeu*educ_middle+land_ceeu*skill_middle+land_ceeu*skill_hi+
                  land_seu*educ_high+land_seu*educ_middle+land_seu*skill_middle+land_seu*skill_hi+
                  land_neu_poor*educ_high+land_neu_poor*educ_middle+land_neu_poor*skill_middle+land_neu_poor*skill_hi

```

```

educ_high+educ_middle+skill_middle+skill_hi,data=data %>% filter(integration_NA==0),se_type="stata

summary(mod9)
#>
#> Call:
#> lm_robust(formula = percent_novotes ~ land_ty * educ_high + land_ty *
#>   educ_middle + land_ty * skill_middle + land_ty * skill_hi +
#>   land_asia * educ_high + land_asia * educ_middle + land_asia *
#>   skill_middle + land_asia * skill_hi + land_ceeu * educ_high +
#>   land_ceceu * educ_middle + land_ceceu * skill_middle + land_ceceu *
#>
#> Standard error type:  HC1
#>
#> Coefficients:
#>
#>             Estimate Std. Error t value Pr(>|t|) CI Lower
#> (Intercept) 23.839    2.049 11.6330 4.624e-29 19.8163
#> land_ty      28.913    2.165 13.3542 5.932e-37 24.6634
#> educ_high     9.674    2.590  3.7344 2.012e-04  4.5891
#> educ_middle    4.939    1.959  2.5213 1.188e-02  1.0940
#> skill_middle   1.233    2.255  0.5470 5.845e-01 -3.1920
#> skill_hi       3.407    1.985  1.7163 8.648e-02 -0.4894
#> land_asia     19.624    2.634  7.4490 2.383e-13 14.4532
#> land_ceceu    18.166    5.078  3.5770 3.678e-04  8.1976
#> land_seu       6.460    2.961  2.1818 2.941e-02  0.6483
#> land_neu_poor  23.828    4.254  5.6013 2.903e-08 15.4780
#> land_ty:educ_high -13.333   3.345 -3.9866 7.300e-05 -19.8980
#> land_ty:educ_middle -5.590   2.176 -2.5686 1.039e-02 -9.8611
#> land_ty:skill_middle -5.804   3.020 -1.9218 5.498e-02 -11.7320
#> land_ty:skill_hi    -5.844   2.210 -2.6450 8.324e-03 -10.1815
#> educ_high:land_asia -4.581   4.694 -0.9760 3.294e-01 -13.7948
#> educ_middle:land_asia -6.434   2.938 -2.1896 2.883e-02 -12.2017
#> skill_middle:land_asia -8.926   5.278 -1.6914 9.115e-02 -19.2855
#> skill_hi:land_asia -6.103   2.941 -2.0751 3.829e-02 -11.8760
#> educ_high:land_ceceu -8.714   6.393 -1.3631 1.732e-01 -21.2620
#> educ_middle:land_ceceu -6.110   5.548 -1.1012 2.711e-01 -17.0004
#> skill_middle:land_ceceu -2.705   5.069 -0.5337 5.937e-01 -12.6561
#> skill_hi:land_ceceu -2.694   5.202 -0.5179 6.047e-01 -12.9050

```

```

#> educ_high:land_seu      -7.705    3.879 -1.9862 4.734e-02 -15.3196
#> educ_middle:land_seu   -2.286    2.776 -0.8236 4.104e-01 -7.7355
#> skill_middle:land_seu  -3.946    3.725 -1.0592 2.898e-01 -11.2573
#> skill_hi:land_seu     -5.451    2.782 -1.9595 5.039e-02 -10.9119
#> educ_high:land_neu_poor -16.340   4.721 -3.4609 5.662e-04 -25.6080
#> educ_middle:land_neu_poor 3.394   4.211  0.8060 4.205e-01 -4.8717
#> skill_middle:land_neu_poor -1.233   3.910 -0.3154 7.525e-01 -8.9085
#> skill_hi:land_neu_poor  -13.407   2.372 -5.6529 2.176e-08 -18.0623
#>                               CI Upper DF
#> (Intercept)            27.861014 821
#> land_ty                33.163016 821
#> educ_high               14.758525 821
#> educ_middle              8.784019 821
#> skill_middle             5.658528 821
#> skill_hi                 7.303447 821
#> land_asia               24.795448 821
#> land_ceeu                28.134251 821
#> land_seu                12.271842 821
#> land_neu_poor            32.178019 821
#> land_ty:educ_high        -6.768364 821
#> land_ty:educ_middle       -1.318090 821
#> land_ty:skill_middle      0.124147 821
#> land_ty:skill_hi          -1.507340 821
#> educ_high:land_asia      4.632530 821
#> educ_middle:land_asia    -0.666363 821
#> skill_middle:land_asia   1.432827 821
#> skill_hi:land_asia       -0.330189 821
#> educ_high:land_ceeu       3.834071 821
#> educ_middle:land_ceeu     4.780450 821
#> skill_middle:land_ceeu   7.245170 821
#> skill_hi:land_ceeu       7.517132 821
#> educ_high:land_seu       -0.090538 821
#> educ_middle:land_seu      3.162607 821
#> skill_middle:land_seu     3.365951 821
#> skill_hi:land_seu         0.009368 821
#> educ_high:land_neu_poor   -7.072926 821
#> educ_middle:land_neu_poor 11.660422 821

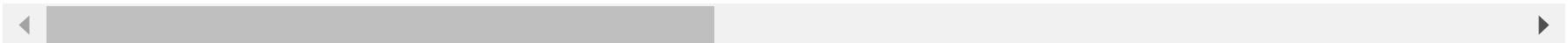
```

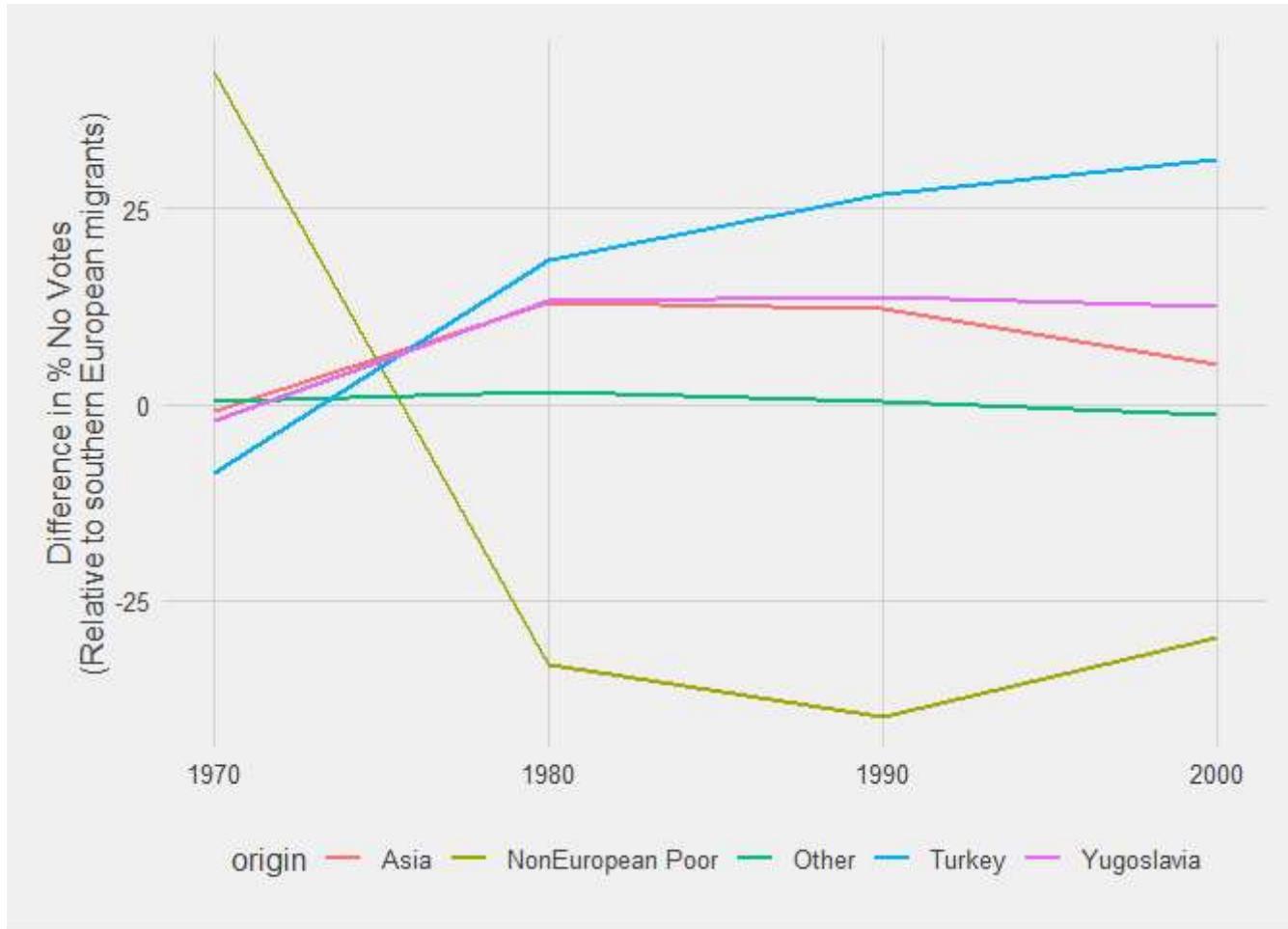


```
}

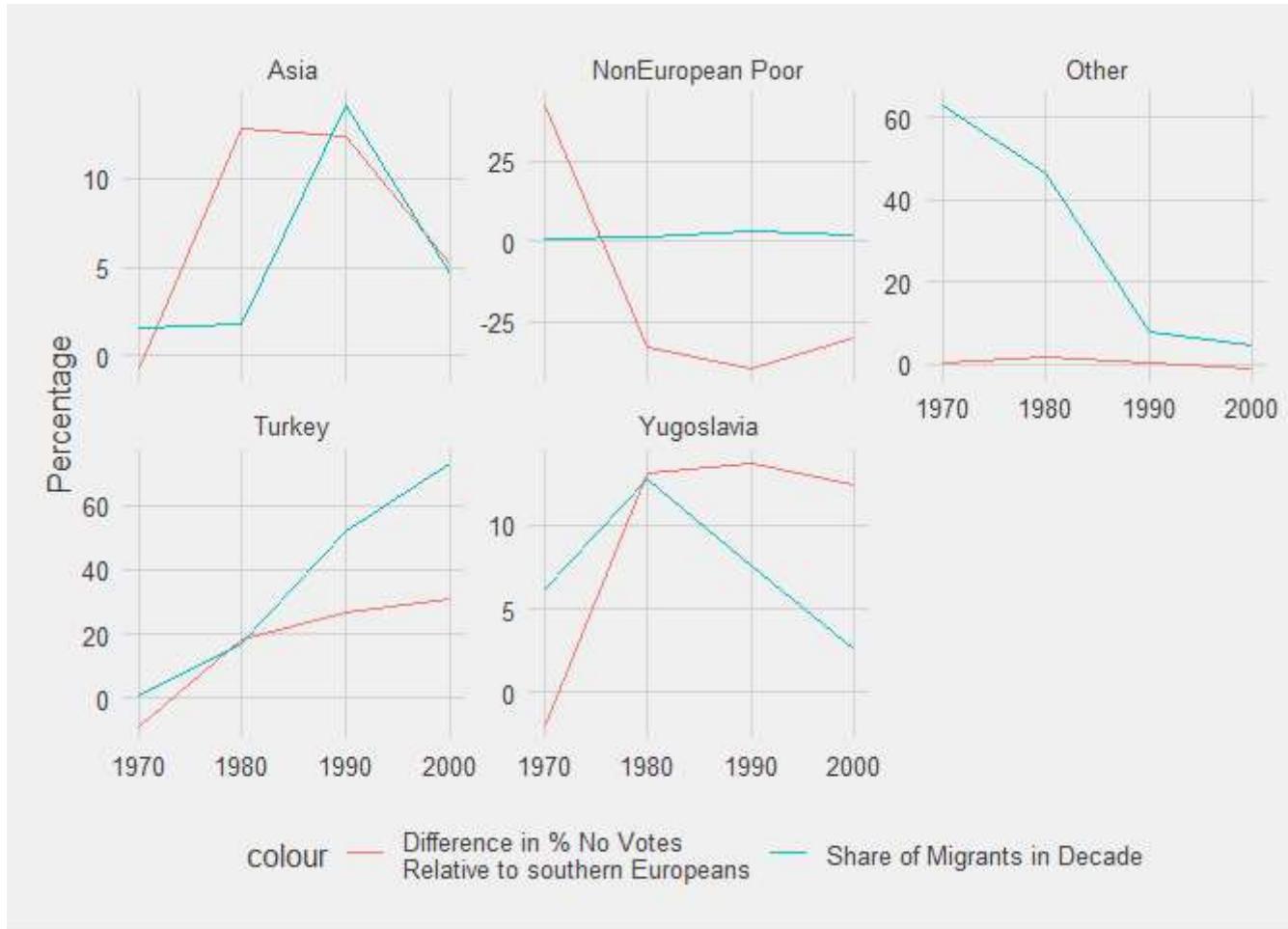
time_fe_fig_data = time_fe_fig_data %>%
  mutate(origin=case_when(
    origin=="ty"~"Turkey",
    origin=="ceeu"~"Yugoslavia",
    origin=="other"~"Other",
    origin=="asia"~"Asia",
    origin=="poor"~"NonEuropean Poor"
  )) %>%
  rowwise() %>%
  mutate(num_migrants=100*nrow(data[data$nationality==origin&data$decade==decade,])/nrow(data[data$decade==decade,]))
  ungroup()

ggplot(time_fe_fig_data,aes(x=decade,y=coef,color=origin)) +
  geom_line(size=1) +
  labs(x="Decade",y="Difference in % No Votes\n(Relative to southern European migrants)") +
  ggthemes::theme_fivethirtyeight() +
  theme(axis.title = element_text(), axis.title.x = element_blank())
```





```
ggplot(time_fe_fig_data,aes(x=decade)) +  
  facet_wrap(~origin,scales="free_y") +  
  geom_line(aes(y=coef,color="Difference in % No Votes\nRelative to southern Europeans"))+  
  geom_line(aes(y=num_migrants,color="Share of Migrants in Decade")) +  
  labs(y="Percentage") +  
  ggthemes::theme_fivethirtyeight() +  
  theme(axis.title = element_text(), axis.title.x = element_blank())
```



```
mod11 = lm_robust(percent_novotes~land_ty*decade+
                    land_ceeu*decade+
                    #land_seu*decade+
                    land_asia*decade+
                    land_neu_poor*decade+
                    land_other*decade+
                    age40+age60+age60plus+
                    male+kids+educ_middle+educ_high+
                    married+skill_hi+skill_middle+born_ch,data=data,se_type="stata")
```

```
summary(mod11)
```

```

#>
#> Call:
#> lm_robust(formula = percent_novotes ~ land_ty * decade + land_ceeu *
#>           decade + land_asia * decade + land_neu_poor * decade + land_other *
#>           decade + age40 + age60 + age60plus + male + kids + educ_middle +
#>           educ_high + married + skill_hi + skill_middle + born_ch,
#>           data = data, se_type = "stata")
#>
#> Standard error type:  HC1
#>
#> Coefficients:
#>
#>             Estimate Std. Error   t value Pr(>|t|) CI Lower
#> (Intercept)  28.53790   2.4760   11.52603 2.143e-29  23.6808
#> land_ty      -9.41792   2.2310  -4.22136 2.589e-05 -13.7945
#> decade1980    0.03776   2.7472   0.01374 9.890e-01  -5.3515
#> decade1990    2.33896   2.4101   0.97046 3.320e-01  -2.3890
#> decade2000   -1.55693   2.4253  -0.64195 5.210e-01  -6.3147
#> land_ceeu     -2.96178   3.0997 -0.95550 3.395e-01  -9.0425
#> land_asia     -0.49251   2.5321 -0.19450 8.458e-01  -5.4598
#> land_neu_poor 42.24447   2.4554  17.20474 3.503e-60  37.4277
#> land_other     0.62292   2.4114   0.25832 7.962e-01  -4.1076
#> age40          1.03758   1.1186   0.92753 3.538e-01  -1.1569
#> age60          0.73322   1.3401   0.54713 5.844e-01  -1.8957
#> age60plus      0.92239   1.9450   0.47423 6.354e-01  -2.8932
#> male            1.35337   0.8412   1.60877 1.079e-01  -0.2969
#> kids            1.06115   1.1597   0.91501 3.603e-01  -1.2139
#> educ_middle    -1.27048   0.8240 -1.54183 1.233e-01  -2.8869
#> educ_high       -0.61585   1.3557 -0.45425 6.497e-01  -3.2754
#> married         0.40151   1.1666   0.34416 7.308e-01  -1.8871
#> skill_hi        -0.39876   0.8539 -0.46698 6.406e-01  -2.0739
#> skill_middle   -2.59041   1.0966 -2.36215 1.831e-02  -4.7417
#> born_ch         -1.43578   0.9576 -1.49932 1.340e-01  -3.3144
#> land_ty:decade1980 18.66765   3.4422   5.42311 6.922e-08  11.9150
#> land_ty:decade1990 26.09829   2.6725   9.76534 8.084e-22  20.8555
#> land_ty:decade2000 30.28478   2.6755  11.31923 1.880e-28  25.0362
#> decade1980:land_ceeu 12.94298   4.4099   2.93499 3.391e-03  4.2921
#> decade1990:land_ceeu 13.69554   3.9097   3.50297 4.750e-04  6.0259

```

```

#> decade2000:land_ceeu      11.85640   5.0185   2.36254 1.829e-02   2.0116
#> decade1980:land_asia    11.40339   3.6896   3.09068 2.037e-03   4.1655
#> decade1990:land_asia    10.33374   2.9381   3.51719 4.505e-04   4.5701
#> decade2000:land_asia     3.85272   3.2969   1.16858 2.428e-01   -2.6149
#> decade1980:land_neu_poor -33.09051   5.8252   -5.68056 1.639e-08   -44.5179
#> decade1990:land_neu_poor -40.86129   3.8677  -10.56476 3.950e-25   -48.4486
#> decade2000:land_neu_poor -31.26670   4.8980   -6.38361 2.366e-10   -40.8751
#> decade1980:land_other     1.01729   3.1376   0.32422 7.458e-01   -5.1378
#> decade1990:land_other     0.27331   3.3114   0.08254 9.342e-01   -6.2226
#> decade2000:land_other     -2.00244   3.7930  -0.52794 5.976e-01   -9.4431
#>
#>                               CI Upper   DF
#> (Intercept)                  33.3950 1362
#> land_ty                     -5.0413 1362
#> decade1980                   5.4270 1362
#> decade1990                   7.0670 1362
#> decade2000                   3.2008 1362
#> land_ceeu                    3.1189 1362
#> land_asia                    4.4748 1362
#> land_neu_poor                 47.0612 1362
#> land_other                    5.3534 1362
#> age40                        3.2320 1362
#> age60                        3.3621 1362
#> age60plus                     4.7380 1362
#> male                          3.0036 1362
#> kids                          3.3362 1362
#> educ_middle                   0.3460 1362
#> educ_high                     2.0437 1362
#> married                       2.6901 1362
#> skill_hi                      1.2764 1362
#> skill_middle                  -0.4391 1362
#> born_ch                       0.4428 1362
#> land_ty:decade1980            25.4203 1362
#> land_ty:decade1990            31.3410 1362
#> land_ty:decade2000            35.5334 1362
#> decade1980:land_ceeu         21.5939 1362
#> decade1990:land_ceeu         21.3652 1362
#> decade2000:land_ceeu         21.7012 1362

```

```

#> decade1980:land_asia      18.6413 1362
#> decade1990:land_asia      16.0974 1362
#> decade2000:land_asia      10.3203 1362
#> decade1980:land_neu_poor -21.6631 1362
#> decade1990:land_neu_poor -33.2740 1362
#> decade2000:land_neu_poor -21.6583 1362
#> decade1980:land_other     7.1724 1362
#> decade1990:land_other     6.7692 1362
#> decade2000:land_other     5.4382 1362
#>
#> Multiple R-squared:  0.3777 ,   Adjusted R-squared:  0.3621
#> F-statistic: 2.744e+12 on 34 and 1362 DF,  p-value: < 2.2e-16

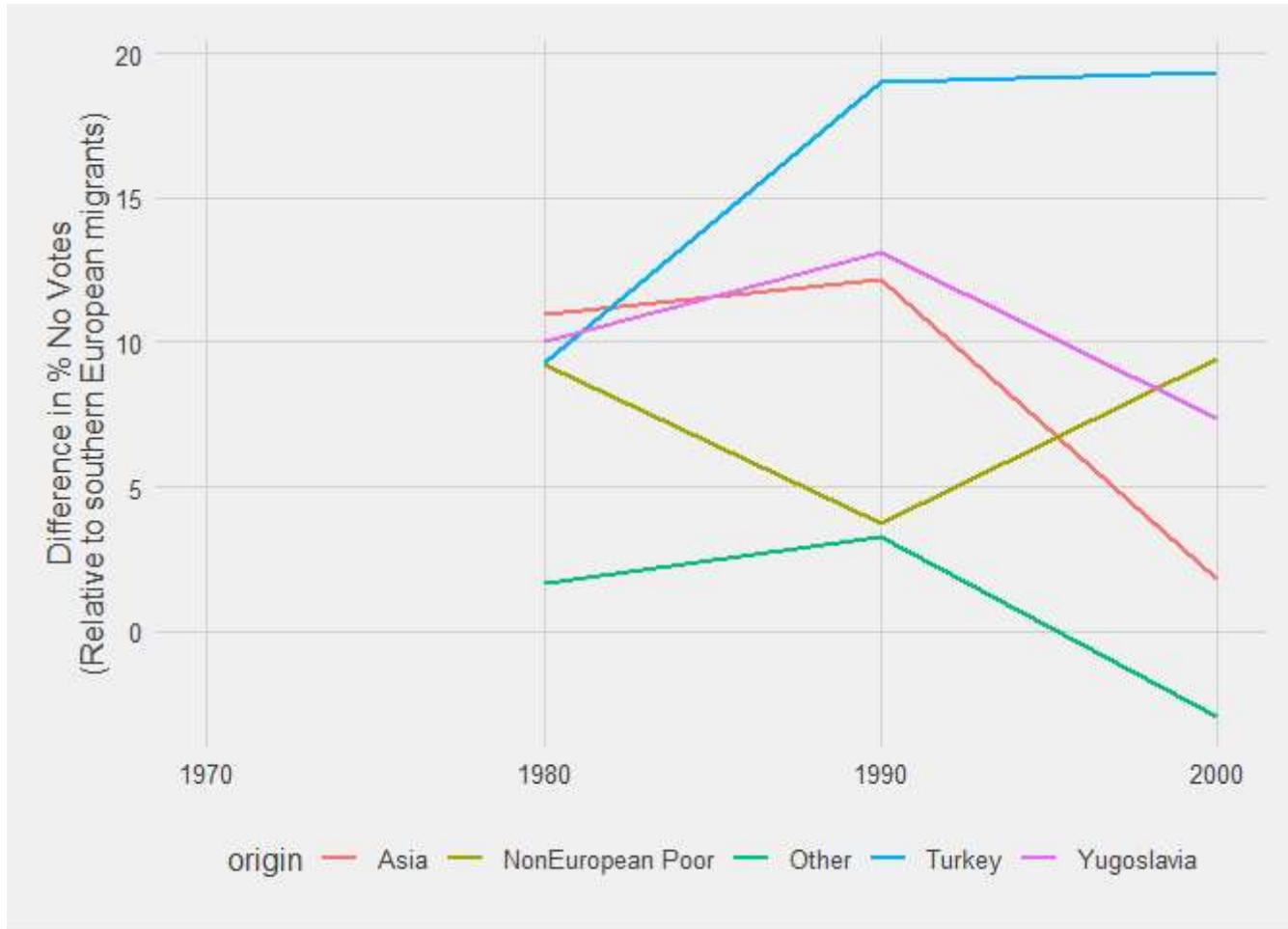
time_fe_results = tidy(mod11)
time_fe_fig_data = data.frame()
for(country in c("ty","ceeu","other","asia","poor")){
  decades=c(1970,1980,1990,2000)
  coeffs = time_fe_results$estimate[grep(country,time_fe_results$term)]
  time_fe_fig_data = bind_rows(time_fe_fig_data,
                                data.frame(decade=decades,
                                           coef=coeffs,
                                           origin=country))
}

time_fe_fig_data = time_fe_fig_data %>%
  rowwise() %>%
  mutate(coef=coef+time_fe_results$estimate[grep(paste0(origin),time_fe_results$term)][1]+time_fe_results$estimate[gr
ungroup() %>%
  mutate(origin=case_when(
    origin=="ty"~"Turkey",
    origin=="ceeu"~"Yugoslavia",
    origin=="other"~"Other",
    origin=="asia"~"Asia",
    origin=="poor"~"NonEuropean Poor"
  )
}

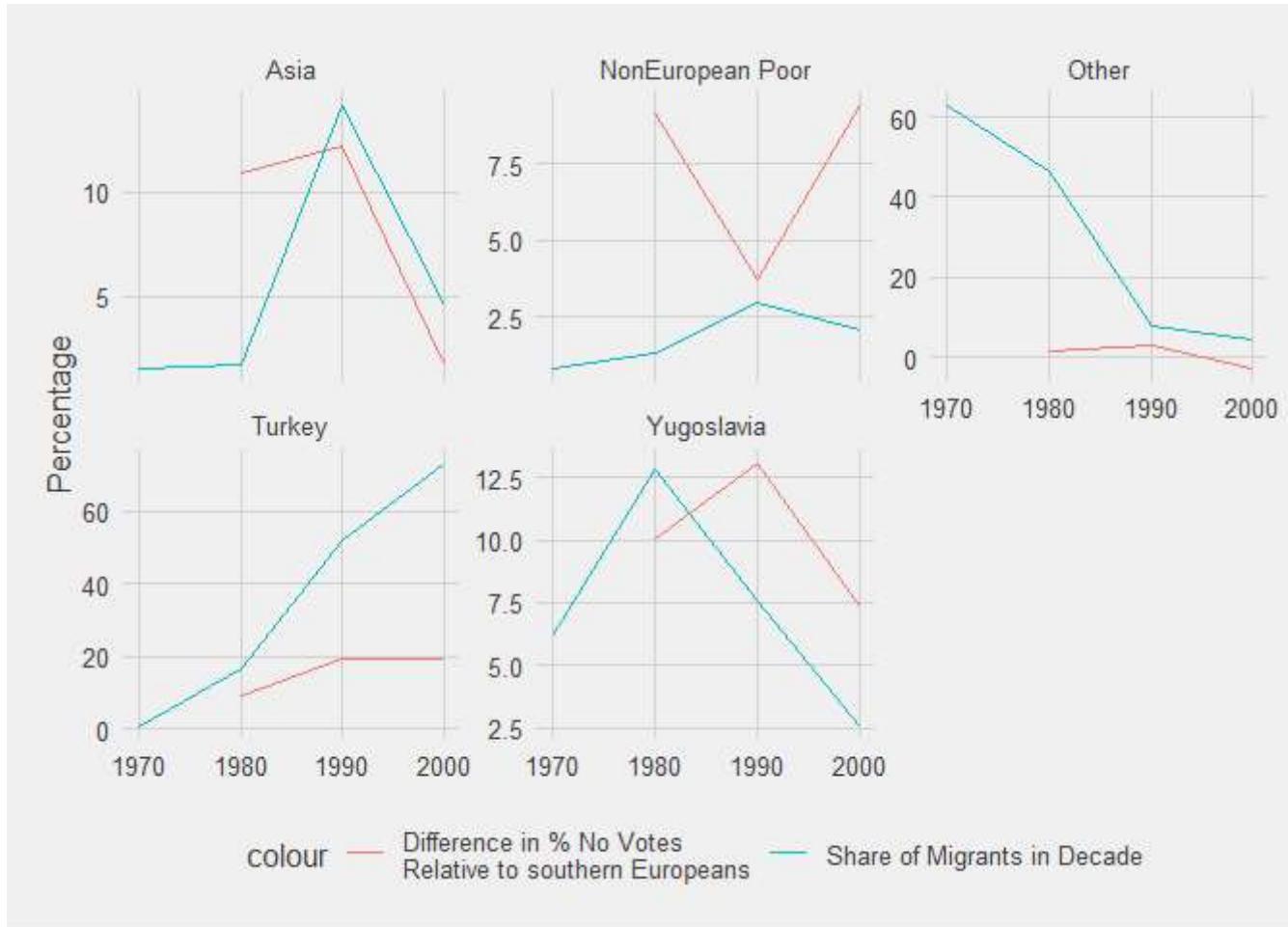
```

```
) %>%  
rowwise() %>%  
mutate(num_migrants=100*nrow(data[data$nationality==origin&data$decade==decade,])/nrow(data[data$decade==decade,]))  
ungroup()  
  
ggplot(time_fe_fig_data,aes(x=decade,y=coef,color=origin)) +  
  geom_line(size=1) +  
  labs(x="Decade",y="Difference in % No Votes\n(Relative to southern European migrants)") +  
  ggthemes::theme_fivethirtyeight() +  
  theme(axis.title = element_text(), axis.title.x = element_blank())  
#> Warning: Removed 5 row(s) containing missing values (geom_path).
```





```
ggplot(time_fe_fig_data,aes(x=decade)) +
  facet_wrap(~origin,scales="free_y") +
  geom_line(aes(y=coef,color="Difference in % No Votes\nRelative to southern Europeans"))+
  geom_line(aes(y=num_migrants,color="Share of Migrants in Decade")) +
  labs(y="Percentage") +
  ggthemes::theme_fivethirtyeight() +
  theme(axis.title = element_text(), axis.title.x = element_blank())
#> Warning: Removed 1 row(s) containing missing values (geom_path).
```



```
mod12 = lm_robust(percent_novotes~land_ty*decade+
                    land_ceeu*decade+
                    #land_seu*decade+
                    land_asia*decade+
                    land_neu_poor*decade+
                    land_other*decade+
                    age40+age60+age60plus+
                    male+kids+educ_middle+educ_high+
                    married+skill_hi+skill_middle+born_ch+
                    lang_good+lang_perfect+factor(integration_integrated)+
```

```

factor(integration_assimilated)+integration_adjusted+
integration_nodifference,data=data,se_type="stata")

summary(mod12)
#> 3 coefficients not defined because the design matrix is rank deficient
#>
#> Call:
#> lm_robust(formula = percent_novotes ~ land_ty * decade + land_ceeu *
#>     decade + land_asia * decade + land_neu_poor * decade + land_other * +
#>     decade + age40 + age60 + age60plus + male + kids + educ_middle +
#>     educ_high + married + skill_hi + skill_middle + born_ch +
#>     lang_good + lang_perfect + factor(integration_integrated) +
#>
#> Standard error type: HC1
#>
#> Coefficients: (3 not defined because the design matrix is rank deficient)
#>
#>             Estimate Std. Error t value Pr(>|t|)

#> (Intercept)      34.44423   3.3199 10.37498 1.326e-22
#> land_ty          -9.94118   2.5999 -3.82369 1.514e-04
#> decade1980       -1.96840   3.3228 -0.59239 5.539e-01
#> decade1990        4.03834   2.6155  1.54400 1.233e-01
#> decade2000       -1.84093   2.5972 -0.70881 4.788e-01
#> land_ceeu         -2.94272   2.5709 -1.14462 2.530e-01
#> land_asia         -5.67711   3.0323 -1.87218 6.188e-02
#> land_neu_poor     10.63740   2.0774  5.12043 4.652e-07
#> land_other        -6.03058   2.6886 -2.24300 2.542e-02
#> age40              1.39273   1.2271  1.13498 2.570e-01
#> age60              0.06761   1.5133  0.04468 9.644e-01
#> age60plus         -2.88474   2.2685 -1.27163 2.042e-01
#> male                1.80368   1.0119  1.78252 7.539e-02
#> kids                -1.46545   1.3818 -1.06057 2.895e-01
#> educ_middle        -0.09853   0.8962 -0.10994 9.125e-01
#> educ_high           1.25377   1.6728  0.74951 4.540e-01
#> married              1.09653   1.4478  0.75736 4.493e-01
#> skill_hi            1.32329   1.0677  1.23934 2.159e-01
#> skill_middle         0.86627   1.4887  0.58190 5.609e-01
#> born_ch             -2.79795   1.1584 -2.41530 1.615e-02

```

```

#> lang_good           -2.36105   2.3292 -1.01367 3.113e-01
#> lang_perfect        -3.21814   2.2547 -1.42729 1.542e-01
#> factor(integration_integrated)1 -4.98452   1.0572 -4.71473 3.298e-06
#> factor(integration_integrated)2 -1.74155   1.6743 -1.04019 2.988e-01
#> factor(integration_assimilated)1 -3.01204   1.6516 -1.82371 6.891e-02
#> factor(integration_assimilated)1.5 -24.52478  2.7220 -9.00998 7.309e-18
#> factor(integration_assimilated)2    0.32405   1.5563  0.20822 8.352e-01
#> integration_adjusted      -3.62465   3.1392 -1.15465 2.489e-01
#> integration_nodifference -4.23264   1.3205 -3.20529 1.452e-03
#> land_ty:decade1980       20.56778   4.3780  4.69797 3.566e-06
#> land_ty:decade1990       27.30098   3.1466  8.67629 9.071e-17
#> land_ty:decade2000       35.06799   3.0291 11.57719 4.317e-27
#> decade1980:land_ceeu     9.90801   8.8106  1.12455 2.614e-01
#> decade1990:land_ceeu     10.25316   4.4783  2.28950 2.255e-02
#> decade2000:land_ceeu     14.11943   3.7868  3.72854 2.190e-04
#> decade1980:land_asia      NA        NA        NA        NA
#> decade1990:land_asia     16.60221   3.6614  4.53441 7.546e-06
#> decade2000:land_asia     13.30817   5.0685  2.62566 8.964e-03
#> decade1980:land_neu_poor    NA        NA        NA        NA
#> decade1990:land_neu_poor    NA        NA        NA        NA
#> decade2000:land_neu_poor   11.85272   2.7381  4.32881 1.876e-05
#> decade1980:land_other      3.83514   4.1262  0.92945 3.532e-01
#> decade1990:land_other      11.20817   4.5457  2.46568 1.407e-02
#> decade2000:land_other      11.50397   6.1498  1.87063 6.209e-02
#>                                     CI Lower CI Upper DF
#> (Intercept)          27.9185  40.9700 420
#> land_ty              -15.0516 -4.8308 420
#> decade1980            -8.4999  4.5631 420
#> decade1990            -1.1028  9.1794 420
#> decade2000            -6.9461  3.2642 420
#> land_ceeu             -7.9962  2.1107 420
#> land_asia             -11.6376  0.2834 420
#> land_neu_poor          6.5539  14.7209 420
#> land_other             -11.3154 -0.7457 420
#> age40                 -1.0193  3.8047 420
#> age60                 -2.9069  3.0422 420
#> age60plus             -7.3438  1.5743 420

```

```

#> male                      -0.1853  3.7926 420
#> kids                       -4.1815  1.2506 420
#> educ_middle                 -1.8600  1.6630 420
#> educ_high                   -2.0343  4.5419 420
#> married                     -1.7494  3.9424 420
#> skill_hi                    -0.7755  3.4221 420
#> skill_middle                 -2.0600  3.7925 420
#> born_ch                     -5.0750  -0.5209 420
#> lang_good                    -6.9394  2.2173 420
#> lang_perfect                  -7.6501  1.2138 420
#> factor(integration_integrated)1 -7.0626  -2.9064 420
#> factor(integration_integrated)2 -5.0325   1.5494 420
#> factor(integration_assimilated)1 -6.2585   0.2344 420
#> factor(integration_assimilated)1.5 -29.8751 -19.1744 420
#> factor(integration_assimilated)2 -2.7350   3.3831 420
#> integration_adjusted          -9.7951   2.5458 420
#> integration_nodifference      -6.8283  -1.6370 420
#> land_ty:decade1980            11.9622  29.1733 420
#> land_ty:decade1990            21.1159  33.4861 420
#> land_ty:decade2000            29.1140  41.0220 420
#> decade1980:land_ceeu         -7.4104  27.2264 420
#> decade1990:land_ceeu          1.4504  19.0559 420
#> decade2000:land_ceeu          6.6759  21.5630 420
#> decade1980:land_asia           NA       NA   NA
#> decade1990:land_asia          9.4053  23.7991 420
#> decade2000:land_asia          3.3454  23.2710 420
#> decade1980:land_neu_poor        NA       NA   NA
#> decade1990:land_neu_poor        NA       NA   NA
#> decade2000:land_neu_poor        6.4706  17.2348 420
#> decade1980:land_other           -4.2755  11.9458 420
#> decade1990:land_other           2.2731  20.1433 420
#> decade2000:land_other           -0.5842  23.5922 420
#>
#> Multiple R-squared:  0.6886 ,     Adjusted R-squared:  0.6589
#> F-statistic:    NA on 40 and 420 DF,  p-value: NA

```

```
time_fe_results = tidy(mod12)
```

```

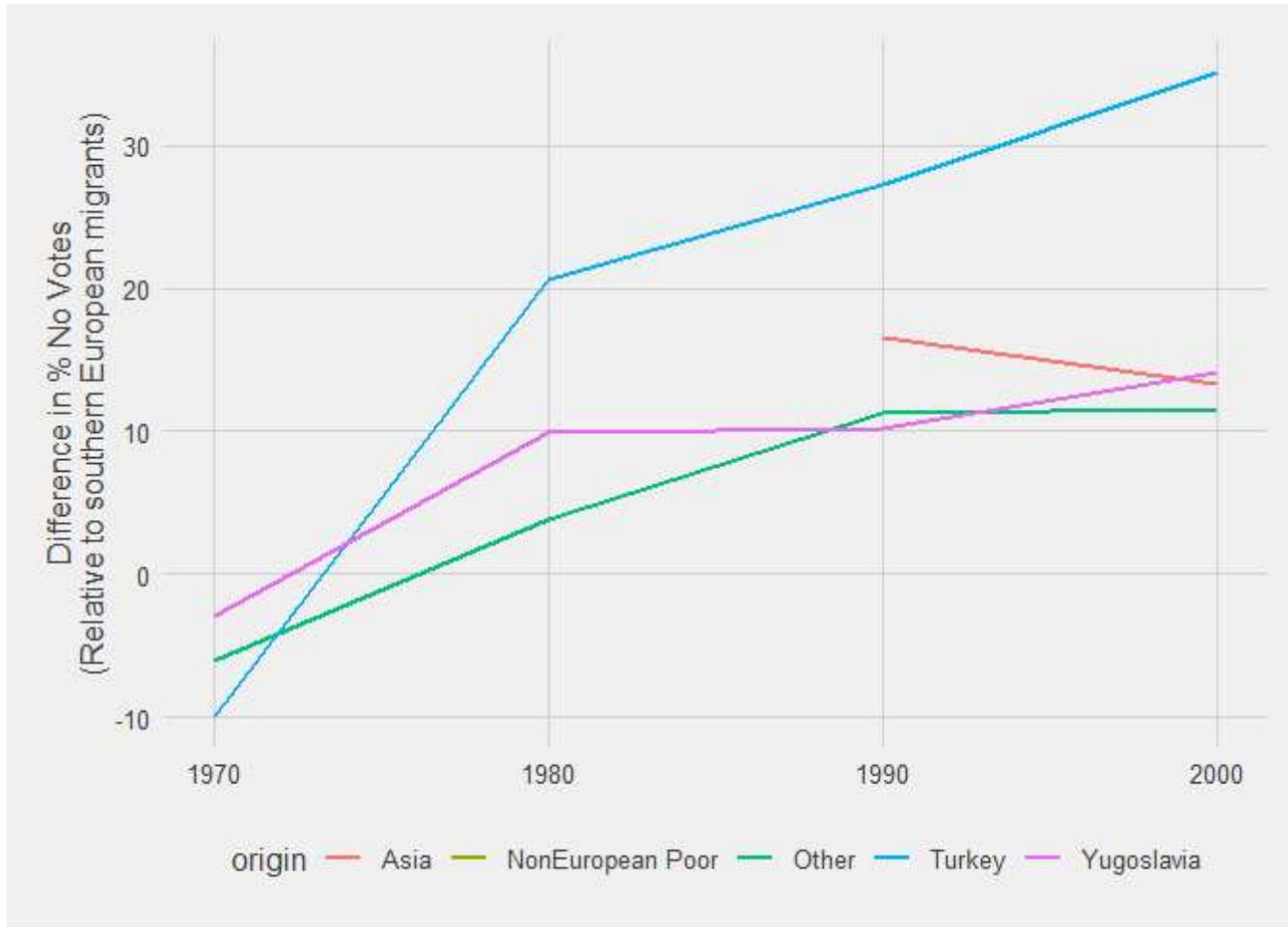
#> 3 coefficients not defined because the design matrix is rank deficient
time_fe_fig_data = data.frame()
for(country in c("ty","ceeu","other","asia","poor")){
  decades=c(1970,1980,1990,2000)
  coeffs = time_fe_results$estimate[grep(country,time_fe_results$term)]
  time_fe_fig_data = bind_rows(time_fe_fig_data,
                                data.frame(decade=decades,
                                           coef=coeffs,
                                           origin=country))
}

time_fe_fig_data = time_fe_fig_data %>%
  mutate(origin=case_when(
    origin=="ty"~"Turkey",
    origin=="ceeu"~"Yugoslavia",
    origin=="other"~"Other",
    origin=="asia"~"Asia",
    origin=="poor"~"NonEuropean Poor"
  )) %>%
  rowwise() %>%
  mutate(num_migrants=100*nrow(data[data$nationality==origin&data$decade==decade,])/nrow(data[data$decade==decade,])))
  ungroup()

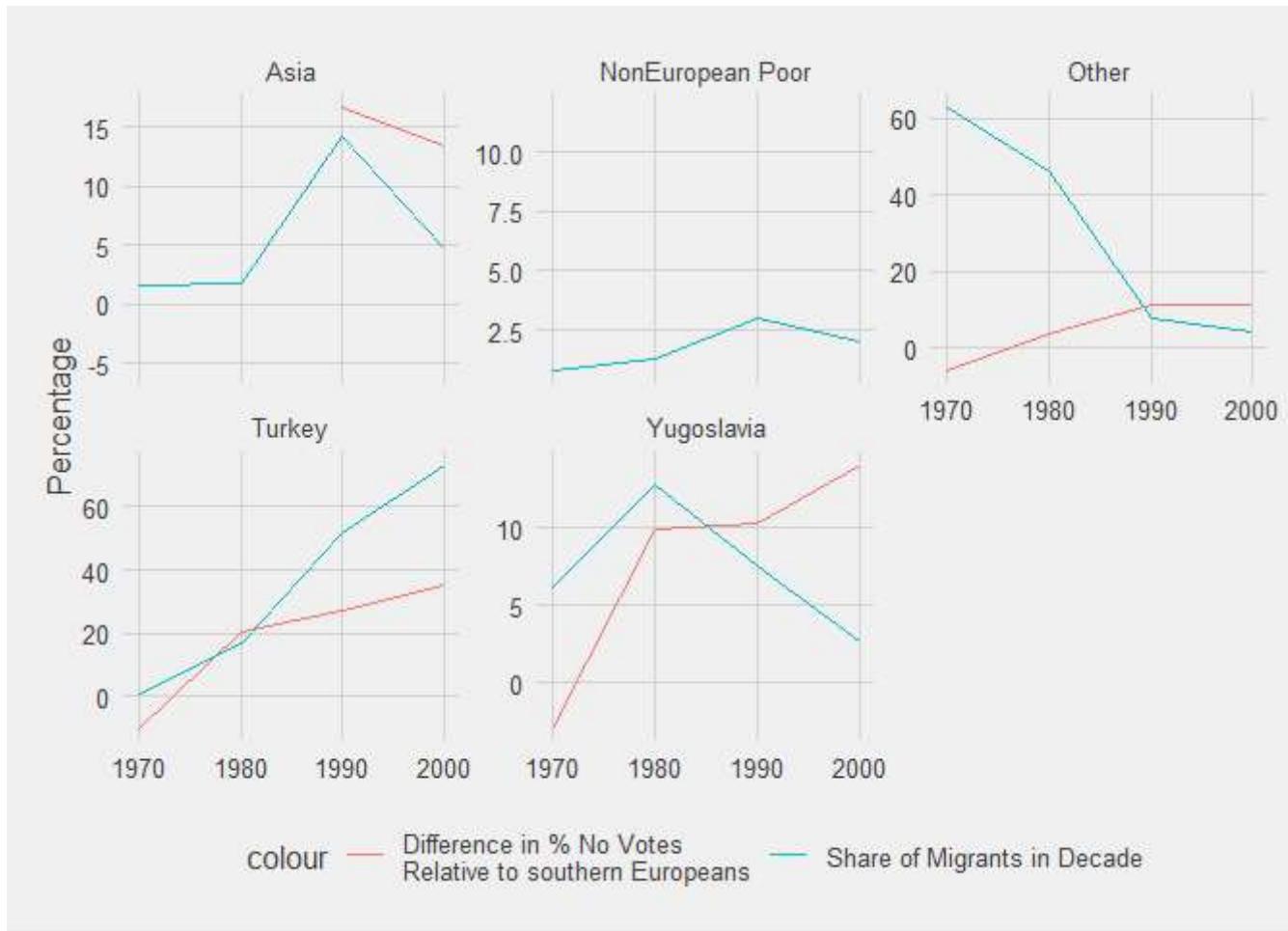
ggplot(time_fe_fig_data,aes(x=decade,y=coef,color=origin)) +
  geom_line(size=1) +
  labs(x="Decade",y="Difference in % No Votes\n(Relative to southern European migrants)") +
  ggthemes::theme_fivethirtyeight() +
  theme(axis.title = element_text(), axis.title.x = element_blank())

```





```
ggplot(time_fe_fig_data,aes(x=decade)) +
  facet_wrap(~origin,scales="free_y") +
  geom_line(aes(y=coef,color="Difference in % No Votes\nRelative to southern Europeans"))+
  geom_line(aes(y=num_migrants,color="Share of Migrants in Decade")) +
  labs(y="Percentage") +
  ggthemes::theme_fivethirtyeight() +
  theme(axis.title = element_text(), axis.title.x = element_blank())
```



```
### Are there any obvious demographic differences
demog_diffs = data %>%
  group_by(nationality) %>%
  summarize_at(vars(year80,year90,year00,
                    age40,age60,age60plus,
                    male,kids,educ_middle,educ_high,
                    married,lang_perfect,lang_good,lang_insufficient,lang_NA,
                    integration_NA,integration_integrated,integration_assimilated,integration_adjusted,integration_no
                    skill_hi,skill_middle,born_ch),mean,na.rm=TRUE) %>%
  ungroup() %>%
  pivot_longer(cols=year80:born_ch,names_to = "variable",values_to = "mean") %>%
```

```

pivot_wider(names_from=nationality,values_from=mean)

# add changes in migration flow share as a control
data = data %>%
  rowwise() %>%
  mutate(num_migrants=100*nrow(data[data$nationality==nationality&data$decade==decade,])/nrow(data)) %>%
  ungroup()

summary(lm_robust(percent_novotes~land_ty*decade+
  land_ceeu*decade+
  #land_seu*decade+
  land_asia*decade+
  land_neu_poor*decade+
  land_other*decade+
  age40+age60+age60plus+
  male+kids+educ_middle+educ_high+
  married+skill_hi+skill_middle+born_ch,data=data,se_type="stata"))

#>
#> Call:
#> lm_robust(formula = percent_novotes ~ land_ty * decade + land_ceeu *
#>   decade + land_asia * decade + land_neu_poor * decade + land_other *
#>   decade + age40 + age60 + age60plus + male + kids + educ_middle +
#>   educ_high + married + skill_hi + skill_middle + born_ch,
#>   data = data, se_type = "stata")
#>
#> Standard error type: HC1
#>
#> Coefficients:
#>             Estimate Std. Error t value Pr(>|t|) CI Lower
#> (Intercept) 28.53790  2.4760 11.52603 2.143e-29 23.6808
#> land_ty      -9.41792  2.2310 -4.22136 2.589e-05 -13.7945
#> decade1980    0.03776  2.7472  0.01374 9.890e-01 -5.3515
#> decade1990    2.33896  2.4101  0.97046 3.320e-01 -2.3890
#> decade2000   -1.55693  2.4253 -0.64195 5.210e-01 -6.3147
#> land_ceeu   -2.96178  3.0997 -0.95550 3.395e-01 -9.0425
#> land_asia    -0.49251  2.5321 -0.19450 8.458e-01 -5.4598
#> land_neu_poor 42.24447  2.4554 17.20474 3.503e-60 37.4277

```

```

#> land_other          0.62292   2.4114   0.25832 7.962e-01 -4.1076
#> age40              1.03758   1.1186   0.92753 3.538e-01 -1.1569
#> age60              0.73322   1.3401   0.54713 5.844e-01 -1.8957
#> age60plus           0.92239   1.9450   0.47423 6.354e-01 -2.8932
#> male                1.35337   0.8412   1.60877 1.079e-01 -0.2969
#> kids                1.06115   1.1597   0.91501 3.603e-01 -1.2139
#> educ_middle          -1.27048  0.8240   -1.54183 1.233e-01 -2.8869
#> educ_high            -0.61585  1.3557   -0.45425 6.497e-01 -3.2754
#> married              0.40151   1.1666   0.34416 7.308e-01 -1.8871
#> skill_hi             -0.39876  0.8539   -0.46698 6.406e-01 -2.0739
#> skill_middle          -2.59041  1.0966   -2.36215 1.831e-02 -4.7417
#> born_ch              -1.43578  0.9576   -1.49932 1.340e-01 -3.3144
#> land_ty:decade1980  18.66765  3.4422   5.42311 6.922e-08 11.9150
#> land_ty:decade1990  26.09829  2.6725   9.76534 8.084e-22 20.8555
#> land_ty:decade2000  30.28478  2.6755   11.31923 1.880e-28 25.0362
#> decade1980:land_ceeu 12.94298  4.4099   2.93499 3.391e-03 4.2921
#> decade1990:land_ceeu 13.69554  3.9097   3.50297 4.750e-04 6.0259
#> decade2000:land_ceeu 11.85640  5.0185   2.36254 1.829e-02 2.0116
#> decade1980:land_asia 11.40339  3.6896   3.09068 2.037e-03 4.1655
#> decade1990:land_asia 10.33374  2.9381   3.51719 4.505e-04 4.5701
#> decade2000:land_asia 3.85272   3.2969   1.16858 2.428e-01 -2.6149
#> decade1980:land_neu_poor -33.09051 5.8252   -5.68056 1.639e-08 -44.5179
#> decade1990:land_neu_poor -40.86129 3.8677   -10.56476 3.950e-25 -48.4486
#> decade2000:land_neu_poor -31.26670 4.8980   -6.38361 2.366e-10 -40.8751
#> decade1980:land_other  1.01729   3.1376   0.32422 7.458e-01 -5.1378
#> decade1990:land_other  0.27331   3.3114   0.08254 9.342e-01 -6.2226
#> decade2000:land_other -2.00244  3.7930   -0.52794 5.976e-01 -9.4431
#>
#> CI Upper DF
#> (Intercept) 33.3950 1362
#> land_ty      -5.0413 1362
#> decade1980   5.4270 1362
#> decade1990   7.0670 1362
#> decade2000   3.2008 1362
#> land_ceeu    3.1189 1362
#> land_asia    4.4748 1362
#> land_neu_poor 47.0612 1362
#> land_other    5.3534 1362

```

```

#> age40           3.2320 1362
#> age60           3.3621 1362
#> age60plus       4.7380 1362
#> male            3.0036 1362
#> kids             3.3362 1362
#> educ_middle     0.3460 1362
#> educ_high        2.0437 1362
#> married          2.6901 1362
#> skill_hi         1.2764 1362
#> skill_middle    -0.4391 1362
#> born_ch          0.4428 1362
#> land_ty:decade1980 25.4203 1362
#> land_ty:decade1990 31.3410 1362
#> land_ty:decade2000 35.5334 1362
#> decade1980:land_ceeu 21.5939 1362
#> decade1990:land_ceeu 21.3652 1362
#> decade2000:land_ceeu 21.7012 1362
#> decade1980:land_asia 18.6413 1362
#> decade1990:land_asia 16.0974 1362
#> decade2000:land_asia 10.3203 1362
#> decade1980:land_neu_poor -21.6631 1362
#> decade1990:land_neu_poor -33.2740 1362
#> decade2000:land_neu_poor -21.6583 1362
#> decade1980:land_other   7.1724 1362
#> decade1990:land_other   6.7692 1362
#> decade2000:land_other   5.4382 1362
#>
#> Multiple R-squared:  0.3777 ,   Adjusted R-squared:  0.3621
#> F-statistic: 2.744e+12 on 34 and 1362 DF,  p-value: < 2.2e-16

# within turkey, what is a good explainer?
summary(lm_robust(percent_novotes~num_migrants+
                    age40+age60+age60plus+
                    male+kids+educ_middle+educ_high+
                    married+skill_hi+skill_middle+born_ch,data=data %>% filter(nationality=="Turkey"),se_type="stata")
#>
#> Call:

```

```

#> lm_robust(formula = percent_novotes ~ num_migrants + age40 +
#>     age60 + age60plus + male + kids + educ_middle + educ_high +
#>     married + skill_hi + skill_middle + born_ch, data = data %>%
#>     filter(nationality == "Turkey"), se_type = "stata")
#>
#> Standard error type: HC1
#>
#> Coefficients:
#>
#>             Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper DF
#> (Intercept) 42.6853    2.36196 18.0720 6.040e-60 38.04772 47.32288 682
#> num_migrants  0.2341    0.07129  3.2843 1.075e-03  0.09416  0.37410 682
#> age40         2.7066    1.51792  1.7831 7.501e-02 -0.27375  5.68695 682
#> age60         3.0792    1.96501  1.5670 1.176e-01 -0.77900  6.93739 682
#> age60plus     11.3613   3.11420  3.6482 2.842e-04  5.24673 17.47587 682
#> male          -0.6482   1.21204 -0.5348 5.930e-01 -3.02793  1.73161 682
#> kids           2.2398    1.93975  1.1547 2.486e-01 -1.56881  6.04837 682
#> educ_middle   -2.5900   1.08075 -2.3965 1.682e-02 -4.71202 -0.46801 682
#> educ_high     -4.3582   2.19086 -1.9893 4.707e-02 -8.65982 -0.05654 682
#> married        -1.2165   1.93354 -0.6291 5.295e-01 -5.01288  2.57993 682
#> skill_hi      -1.4956   1.27069 -1.1770 2.396e-01 -3.99053  0.99934 682
#> skill_middle  -4.5893   1.79744 -2.5533 1.089e-02 -8.11849 -1.06014 682
#> born_ch        2.8259    1.30722  2.1617 3.099e-02  0.25921  5.39252 682
#>
#> Multiple R-squared:  0.06329 , Adjusted R-squared:  0.0468
#> F-statistic: 4.635 on 12 and 682 DF, p-value: 2.892e-07

```

```

# Logit model
data_logit = data
table(data_logit$percent_approved)
#>
#> 0   1
#> 440 957
table(data_logit$nationality,data_logit$percent_approved)
#>
#>          0   1

```

```

#>   Asia          10  92
#> NonEuropean Poor    6  24
#> Other           9 240
#> Southern Europe   13 219
#> Turkey          378 317
#> Yugoslavia       24  65

data_logit <- filter(data_logit, integration_assimilated!=1.5)

set.seed(1)
options(scipen=999)

sample <- sample(c(TRUE, FALSE), nrow(data_logit), replace=TRUE, prob=c(0.9,0.1))
train <- data_logit[sample, ]
test <- data_logit[!sample, ]
mod20 <- glm(percent_approved~land_ty+land_asia+
               land_ceeu+land_seu+land_neu_poor+
               educ_high+educ_middle+skill_middle+skill_hi, family="binomial", data=train)

summary(mod20)
#>
#> Call:
#> glm(formula = percent_approved ~ land_ty + land_asia + land_ceeu +
#>       land_seu + land_neu_poor + educ_high + educ_middle + skill_middle +
#>       skill_hi, family = "binomial", data = train)
#>
#> Deviance Residuals:
#>      Min        1Q     Median        3Q       Max
#> -2.5578  -0.9340   0.3087   0.5282   1.5376
#>
#> Coefficients:
#>             Estimate Std. Error z value     Pr(>|z|)
#> (Intercept)  2.7237    0.4908   5.550 0.0000000285699066 ***
#> land_ty      -3.5397    0.4714  -7.509 0.000000000000594 ***
#> land_asia    -0.8246    0.6353  -1.298     0.194289
#> land_ceeu    -2.3951    0.5622  -4.260 0.000204139390055 ***
#> land_seu     -0.4493    0.5882  -0.764     0.445024

```

```

#> land_neu_poor -2.7238    0.7992   -3.408      0.000654 ***
#> educ_high      0.2589    0.3873   0.669      0.503755
#> educ_middle    0.2964    0.2039   1.454      0.146014
#> skill_middle   0.6793    0.3399   1.998      0.045685 *
#> skill_hi       0.2123    0.2029   1.046      0.295361
#> ---
#> Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#>
#> (Dispersion parameter for binomial family taken to be 1)
#>
#> Null deviance: 966.85 on 738 degrees of freedom
#> Residual deviance: 696.44 on 729 degrees of freedom
#> AIC: 716.44
#>
#> Number of Fisher Scoring iterations: 5
exp(coef(mod20))
#> (Intercept)      land_ty      land_asia      land_ceeu      land_seu
#> 15.23724902    0.02902255   0.43839468   0.09116019   0.63810576
#> land_neu_poor   educ_high    educ_middle   skill_middle   skill_hi
#> 0.06562273     1.29553063   1.34495970   1.97240916   1.23657522

mod21 <- glm(percent_approved~land_ty+land_asia+land_ceeu+land_seu+land_neu_poor+year80+year90+year00+male+kids+marri
summary(mod21)
#>
#> Call:
#> glm(formula = percent_approved ~ land_ty + land_asia + land_ceeu +
#>      land_seu + land_neu_poor + year80 + year90 + year00 + male +
#>      kids + married + born_ch + age60 + educ_high + skill_hi +
#>      lang_good + lang_perfect, family = "binomial", data = train)
#>
#> Deviance Residuals:
#>      Min        1Q      Median        3Q        Max
#> -2.41723  -0.81453   0.06535   0.58691   1.85980
#>
#> Coefficients:

```

```

#>             Estimate Std. Error z value Pr(>|z|)
#> (Intercept) 19.1039   643.9788   0.030  0.97633
#> land_ty     -3.3630    1.1703  -2.874  0.00406 **
#> land_asia   -1.0446    1.2623  -0.828  0.40792
#> land_ceeu   -1.5732    1.2295  -1.280  0.20071
#> land_seu    -0.4780    1.2486  -0.383  0.70186
#> land_neu_poor -2.4942   1.4374  -1.735  0.08270 .
#> year80      -12.6348   643.9783 -0.020  0.98435
#> year90      -15.5865   643.9775 -0.024  0.98069
#> year00      -16.0574   643.9775 -0.025  0.98011
#> male         -0.1043    0.3180  -0.328  0.74285
#> kids          -0.3850    0.4772  -0.807  0.41978
#> married       0.1005    0.4262  0.236  0.81358
#> born_ch      0.6226    0.3458  1.800  0.07180 .
#> age60         0.2682    0.3923  0.684  0.49416
#> educ_high    -0.6751    0.5502  -1.227  0.21984
#> skill_hi     -0.2878    0.3356  -0.858  0.39108
#> lang_good    -0.8092    0.8167  -0.991  0.32173
#> lang_perfect -0.5081    0.7837  -0.648  0.51678
#> ---
#> Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#>
#> (Dispersion parameter for binomial family taken to be 1)
#>
#> Null deviance: 538.65 on 397 degrees of freedom
#> Residual deviance: 358.59 on 380 degrees of freedom
#> (341 observations deleted due to missingness)
#> AIC: 394.59
#>
#> Number of Fisher Scoring iterations: 16
exp(coef(mod21))
#>             (Intercept)           land_ty           land_asia
#> 198027879.4179713129997        0.0346316004330        0.3518230225676
#>           land_ceeu           land_seu           land_neu_poor
#> 0.2073710861235        0.6200326581672        0.0825663134871
#>           year80           year90           year00
#> 0.0000032566845        0.0000001701684        0.0000001062545

```

```

#>           male            kids            married
#> 0.9009163571657 0.6804568085812 1.1057203150321
#>          born_ch        age60      educ_high
#> 1.8637342437382 1.3076244144167 0.5090989778913
#>          skill_hi     lang_good    lang_perfect
#> 0.7499063748159 0.4451996573474 0.6016416011198

mod22 <- glm(percent_approved~land_ty+land_asia+
               land_ceeu+land_seu+land_neu_poor+year80+
               year90+year00+male+kids+married+born_ch+
               age60+age60plus+educ_high+skill_hi+
               lang_good+lang_perfect+factor(integration_integrated)+
               factor(integration_assimilated)+integration_adjusted+
               integration_nodifference, family="binomial", data=train)
#> Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

summary(mod22)
#>
#> Call:
#> glm(formula = percent_approved ~ land_ty + land_asia + land_ceeu +
#>       land_seu + land_neu_poor + year80 + year90 + year00 + male +
#>       kids + married + born_ch + age60 + age60plus + educ_high +
#>       skill_hi + lang_good + lang_perfect + factor(integration_integrated) +
#>       factor(integration_assimilated) + integration_adjusted +
#>       integration_nodifference, family = "binomial", data = train)
#>
#> Deviance Residuals:
#>      Min      1Q   Median      3Q      Max
#> -2.1668 -0.7813  0.0182  0.5542  2.0760
#>
#> Coefficients:
#>                               Estimate Std. Error z value Pr(>|z|)
#> (Intercept)                19.06868 1017.93929  0.019  0.98505
#> land_ty                   -3.66580   1.18485 -3.094  0.00198 **
#> land_asia                  -1.25336   1.28944 -0.972  0.33104
#> land_ceeu                  -1.48188   1.25120 -1.184  0.23627

```

```

#> land_seu                      -0.52242   1.25821  -0.415   0.67799
#> land_neu_poor                  -2.95818   1.46717  -2.016   0.04377 *
#> year80                         -12.98813 1017.93894 -0.013   0.98982
#> year90                         -16.23605 1017.93838 -0.016   0.98727
#> year00                         -16.42789 1017.93837 -0.016   0.98712
#> male                            -0.08098   0.33042  -0.245   0.80640
#> kids                            -0.32257   0.50736  -0.636   0.52492
#> married                         0.12498   0.44045  0.284   0.77660
#> born_ch                         0.74593   0.37597  1.984   0.04726 *
#> age60                           0.36059   0.41089  0.878   0.38017
#> age60plus                       15.58740 2241.48679  0.007   0.99445
#> educ_high                        -0.84465   0.56697  -1.490   0.13628
#> skill_hi                         -0.26589   0.34780  -0.764   0.44458
#> lang_good                        -0.89848   0.85954  -1.045   0.29588
#> lang_perfect                     -0.54321   0.82523  -0.658   0.51037
#> factor(integration_integrated)1  1.15300   0.36450  3.163   0.00156 **
#> factor(integration_integrated)2  0.13838   0.60592  0.228   0.81934
#> factor(integration_assimilated)1 1.09290   0.49722  2.198   0.02795 *
#> factor(integration_assimilated)2 0.30516   0.64435  0.474   0.63578
#> integration_adjusted             17.80284 6522.63863  0.003   0.99782
#> integration_nodifference        1.41263   0.50650  2.789   0.00529 **
#> ---
#> Signif. codes:  0 '****' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#>
#> (Dispersion parameter for binomial family taken to be 1)
#>
#> Null deviance: 538.65 on 397 degrees of freedom
#> Residual deviance: 341.32 on 373 degrees of freedom
#> (341 observations deleted due to missingness)
#> AIC: 391.32
#>
#> Number of Fisher Scoring iterations: 17
exp(coef(mod22))
#>                               (Intercept)                         land_ty
#> 191170344.67824137210846          0.02558379285372
#>                               land_asia                         land_ceeu
#> 0.28554272809128          0.22721120882409

```

```

#>           land_seu                  land_neu_poor
#> 0.59308097605053                 0.05191311384137
#>           year80                  year90
#> 0.00000228732622                 0.0000008887336
#>           year00                  male
#> 0.0000007335967                 0.92221445779237
#>           kids                   married
#> 0.72428339217290                1.13312487617471
#>           born_ch                 age60
#> 2.10839893066960                1.43417000782705
#>           age60plus               educ_high
#> 5881946.09052091185004            0.42970705631973
#>           skill_hi                lang_good
#> 0.76652621869008                0.40718801483329
#>           lang_perfect   factor(integration_integrated)1
#> 0.58087793273999                3.16766952783873
#> factor(integration_integrated)2 factor(integration_assimilated)1
#> 1.14841692629936                2.98291037717863
#> factor(integration_assimilated)2           integration_adjusted
#> 1.35684783457297                53910739.10873291641474
#>           integration_nodifference
#> 4.10674068969319

# Adding ANOVA analysis for Model 7
mod71 = lm(percent_novotes~land_ty+land_asia+
            land_ceeu+land_seu+land_neu_poor+year80+
            year90+year00+male+kids+married+born_ch+
            age60+age60plus+educ_high+skill_hi+
            lang_good+lang_perfect+factor(integration_integrated)+
            factor(integration_assimilated)+integration_adjusted+
            integration_nodifference,data=data %>% filter(integration_NA==0))

summary(mod71)
#>
#> Call:
#> lm(formula = percent_novotes ~ land_ty + land_asia + land_ceeu +
#>     land_seu + land_neu_poor + year80 + year90 + year00 + male +
#>     kids + married + born_ch + age60 + age60plus + educ_high +

```

```

#> skill_hi + lang_good + lang_perfect + factor(integration_integrated) +
#> factor(integration_assimilated) + integration_adjusted +
#> integration_nodifference, data = data %>% filter(integration_NA ==
#> 0))
#>
#> Residuals:
#>    Min     1Q Median     3Q    Max
#> -37.015 -4.788  0.510  4.955 22.187
#>
#> Coefficients:
#>                               Estimate Std. Error t value
#> (Intercept)                  33.7152   3.4694  9.718
#> land_ty                      17.3411   1.8432  9.408
#> land_asia                     6.6341   2.3347  2.842
#> land_ceeu                     5.7281   2.3476  2.440
#> land_seu                      -1.4611   1.8245 -0.801
#> land_neu_poor                 13.5970   3.6893  3.685
#> year80                        0.8469   2.0061  0.422
#> year90                        9.8630   1.9505  5.057
#> year00                        9.4053   1.9991  4.705
#> male                           1.1318   0.9611  1.178
#> kids                           -2.0884   1.4125 -1.479
#> married                        1.5869   1.2935  1.227
#> born_ch                        -3.5065   1.1147 -3.146
#> age60                          -1.3539   1.0867 -1.246
#> age60plus                     -4.6952   2.9882 -1.571
#> educ_high                      2.0588   1.4727  1.398
#> skill_hi                       0.1668   0.9633  0.173
#> lang_good                      -3.9988   2.7156 -1.473
#> lang_perfect                   -5.2630   2.5913 -2.031
#> factor(integration_integrated)1 -5.1211   1.1241 -4.556
#> factor(integration_integrated)2 -1.2908   1.7191 -0.751
#> factor(integration_assimilated)1 -2.8079   1.4170 -1.982
#> factor(integration_assimilated)1.5 -24.1165  2.2115 -10.905
#> factor(integration_assimilated)2    0.5460   1.5627  0.349
#> integration_adjusted            -4.8069   6.0116 -0.800
#> integration_nodifference       -3.4525   1.4448 -2.390

```

```

#>                                     Pr(>|t|)
#> (Intercept)                      < 0.0000000000000002 ***
#> land_ty                           < 0.0000000000000002 ***
#> land_asia                         0.004700 **
#> land_ceeu                          0.015085 *
#> land_seu                           0.423645
#> land_neu_poor                     0.000257 ***
#> year80                            0.673100
#> year90                            0.00000063 ***
#> year00                            0.00000342 ***
#> male                               0.239602
#> kids                               0.139995
#> married                            0.220537
#> born_ch                            0.001770 **
#> age60                             0.213473
#> age60plus                         0.116854
#> educ_high                          0.162825
#> skill_hi                           0.862592
#> lang_good                          0.141604
#> lang_perfect                       0.042861 *
#> factor(integration_integrated)1   0.00000679 ***
#> factor(integration_integrated)2   0.453144
#> factor(integration_assimilated)1  0.048161 *
#> factor(integration_assimilated)1.5 < 0.0000000000000002 ***
#> factor(integration_assimilated)2   0.726972
#> integration_adjusted               0.424372
#> integration_nodifference          0.017289 *
#> ---
#> Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#>
#> Residual standard error: 8.287 on 435 degrees of freedom
#>   (390 observations deleted due to missingness)
#> Multiple R-squared:  0.6528, Adjusted R-squared:  0.6329
#> F-statistic: 32.72 on 25 and 435 DF,  p-value: < 0.000000000000022
anova(mod71)
#> Analysis of Variance Table
#>

```

```

#> Response: percent_novotes
#>
#>   Df Sum Sq Mean Sq F value
#> land_ty           1 36287 36287 528.4192
#> land_asia         1 3340  3340  48.6360
#> land_ceeu         1  975   975  14.2026
#> land_seu          1     1     1  0.0074
#> land_neu_poor    1 2315  2315  33.7125
#> year80            1  543   543  7.9091
#> year90            1  188   188  2.7323
#> year00            1 1639  1639  23.8668
#> male              1  159   159  2.3145
#> kids               1   33    33  0.4838
#> married            1  252   252  3.6640
#> born_ch            1  427   427  6.2157
#> age60              1   31    31  0.4443
#> age60plus          1  181   181  2.6301
#> educ_high          1   53    53  0.7705
#> skill_hi           1   16    16  0.2292
#> lang_good           1   23    23  0.3362
#> lang_perfect        1  456   456  6.6404
#> factor(integration_integrated) 2  580   290  4.2249
#> factor(integration_assimilated) 3 8250  2750 40.0481
#> integration_adjusted      1   28    28  0.4062
#> integration_nodifference 1  392   392  5.7105
#> Residuals          435 29872   69
#>
#>   Pr(>F)
#> land_ty            < 0.0000000000000022 ***
#> land_asia          0.0000000001152 ***
#> land_ceeu          0.0001868 ***
#> land_seu           0.9314018
#> land_neu_poor     0.0000001233821 ***
#> year80             0.0051414 **
#> year90             0.0990570 .
#> year00             0.0000145351006 ***
#> male               0.1289027
#> kids               0.4870951
#> married            0.0562547 .

```

```

#> born_ch          0.0130326 *
#> age60            0.5054312
#> age60plus        0.1055827
#> educ_high         0.3805514
#> skill_hi          0.6323720
#> lang_good          0.5623376
#> lang_perfect       0.0102979 *
#> factor(integration_integrated)    0.0152319 *
#> factor(integration_assimilated) < 0.0000000000000022 ***
#> integration_adjusted      0.5242269
#> integration_nodifference   0.0172885 *
#> Residuals
#> ---
#> Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```

##### Plan for memo #####
# 1. Brief history of immigration to Switzerland
# Show graph of applications over time broken down by country
fig1_data = data %>%
  group_by(decade,nationality) %>%
  summarize(num=n()) %>%
  ungroup() %>%
  rename(origin=nationality)
#> `summarise()` has grouped output by 'decade'. You can override using the
#> ` `.groups` argument.

fig1 = ggplot(fig1_data,aes(x=decade,y=num,fill=origin,group=origin)) +
  geom_area() +
  ggthemes::theme_fivethirtyeight() +
  theme(axis.title = element_text(), axis.title.x = element_blank()) +
  labs(x="Decade",y="Number of immigrant applications",caption="Source: Hainmueller and Hangartner")

ggsave(paste0(charts_folder,"figure1.png"),fig1,width=10,height=5,units="in",dpi=320)

# 2. How do immigrants differ?
table1_data = data %>%

```

```

mutate(percent_novotes=percent_novotes/100) %>%
group_by(nationality) %>%
summarize_at(vars(percent_novotes,percent_approved,age40,age60,age60plus,
male,kids,educ_middle,educ_high,
married,lang_perfect,lang_good,lang_insufficient,lang_NA,
integration_NA,integration_adjusted,integration_nodifference,
skill_hi,skill_middle,born_ch),mean,na.rm=TRUE) %>%
ungroup() %>%
pivot_longer(cols=percent_novotes:born_ch,names_to = "variable",values_to = "mean") %>%
pivot_wider(names_from=nationality,values_from=mean) %>%
mutate(variable=case_when(
  variable=="percent_novotes"~"Average %No votes received",
  variable=="percent_approved"~"Percent approved",
  variable=="age40"~"Percent between age 20 and 40",
  variable=="age60"~"Percent between age 40 and 60",
  variable=="age60plus"~"Percent above age 60",
  variable=="male"~"Percent male",
  variable=="kids"~"Percent with children",
  variable=="educ_middle"~"Percent with high-school education",
  variable=="educ_high"~"Percent with college education",
  variable=="married"~"Percent married",
  variable=="lang_perfect"~"Percent language proficiency close to perfect",
  variable=="lang_good"~"Percent language proficiency is good",
  variable=="lang_insufficient"~"Percent language proficiency insufficient",
  variable=="lang_NA"~"Percent language data missing",
  variable=="integration_NA"~"Percent integration data missing",
  variable=="integration_adjusted"~"Percent that have adjusted to Swiss culture",
  variable=="integration_nodifference"~"Percent who are well assimilated",
  variable=="skill_hi"~"Percent in high-skilled occupation",
  variable=="skill_middle"~"Percent in medium-skilled occupation",
  variable=="born_ch"~"Percent born in Switzerland"
))
library(purrr)
table1_data1 <- table1_data %>%
  mutate(row_num=row_number())
table1_data1 = table1_data1 %>%

```

```

rowwise() %>%
  mutate(min_val=which.min(table1_data1[table1_data1$row_num==row_num,2:7]),
        max_val=which.max(table1_data1[table1_data1$row_num==row_num,2:7])) %>%
  ungroup() %>%
  mutate_all(as.character)
for(i in 1:nrow(table1_data1)) {

  table1_data1[i,as.numeric(table1_data1$min_val[i])+1] = "#00A7E1"
  table1_data1[i,as.numeric(table1_data1$max_val[i])+1] = "#FFA630"
  table1_data1[i,-c(1,8,9,10,as.numeric(table1_data1$min_val[i])+1,as.numeric(table1_data1$max_val[i])+1)] = "#FFFFFF"

}

table1 = reactable(table1_data,
                   defaultPageSize = 200,
                   theme=fivethirtyeight(),
                   columns = list(variable=colDef(name="Variable",align = "center",format=colFormat(percent=TRUE,digits=2)),
                                  Asia=colDef(name="Asia",align = "center",format=colFormat(percent=TRUE,digits=2),centered=TRUE),
                                  `NonEuropean Poor`=colDef(name="Non-European Developing",align = "center",format=colFormat(percent=TRUE,digits=2)),
                                  Other=colDef(name="Other regions",align = "center",format=colFormat(percent=TRUE,digits=2)),
                                  `Southern Europe`=colDef(name="Southern Europe",align = "center",format=colFormat(percent=TRUE,digits=2)),
                                  Turkey=colDef(name="Turkey",align = "center",format=colFormat(percent=TRUE,digits=2)),
                                  Yugoslavia=colDef(name="Yugoslavia",align = "center",format=colFormat(percent=TRUE,digits=2)))

data = data %>% mutate(percent_approved=percent_approved*100)
mod1 = lm_robust(percent_approved~land_ty+land_asia+land_ceeu+land_other+land_neu_poor,data=data,se_type="stata")
mod2 = lm_robust(percent_approved~land_ty+land_asia+land_ceeu+land_other+land_neu_poor+
                  year80+year90+year00+age40+age60+age60plus+
                  male+kids+married+born_ch,data=data,se_type="stata")
mod3 = lm_robust(percent_approved~land_ty+land_asia+land_ceeu+land_other+land_neu_poor+
                  educ_middle+educ_high+skill_hi+skill_middle,data=data,se_type="stata")
mod4 = lm_robust(percent_approved~land_ty+land_asia+land_ceeu+land_other+land_neu_poor+
                  lang_perfect+lang_good+lang_insufficient,data=data,se_type="stata")
mod5 = lm_robust(percent_approved~land_ty+land_asia+land_ceeu+land_other+land_neu_poor+
                  factor(integration_integrated)+factor(integration_assimilated)+)

```

```

        integration_adjusted+integration_nodifference,data=data,se_type="stata")
mod6 = lm_robust(percent_approved~land_ty+land_asia+land_ceeu+land_other+land_neu_poor+
                  year80+year90+year00+age40+age60+age60plus+
                  male+kids+married+born_ch+
                  educ_middle+educ_high+skill_hi+skill_middle+
                  lang_perfect+lang_good+lang_insufficient+
                  factor(integration_integrated)+factor(integration_assimilated)+
                  integration_adjusted+integration_nodifference,data=data,se_type="stata")
mod7 = lm_robust(percent_approved~factor(nationality)*factor(decade),data=data,se_type="stata")
#> Warning in sqrt(diag(vcov_fit$Vcov_hat)): NaNs produced
mod8 = lm_robust(percent_approved~factor(nationality)*factor(decade)+
                  age40+age60+age60plus+
                  male+kids+married+born_ch+
                  educ_middle+educ_high+skill_hi+skill_middle,data=data,se_type="stata")
mod9 = lm_robust(percent_approved~land_ty*factor(decade)+
                  land_ty*age40+land_ty*age60+land_ty*age60plus+
                  land_ty*male+land_ty*kids+land_ty*married+land_ty*born_ch+
                  land_ty*educ_middle+land_ty*educ_high+land_ty*skill_hi+land_ty*skill_middle,data=data,se_type="sta
htmlreg(list(mod1,mod2,mod3,
            mod4,mod5,mod6,mod7,mod8,mod9),file=paste0(results_folder,"regressions.html"),include.ci=FALSE)
#> The table was written to the file 'C:/Users/16094/Documents/Year 1 - Princeton/Fall 2022/Quant Analysis/Project/Re

mod12 = lm_robust(percent_approved~land_ty*decade+
                  land_ceeu*decade+
                  #land_seu*decade+
                  land_asia*decade+
                  land_neu_poor*decade+
                  land_other*decade+
                  age40+age60+age60plus+
                  male+kids+educ_middle+educ_high+
                  married+skill_hi+skill_middle+born_ch,data=data,se_type="stata")

summary(mod12)
#>
#> Call:
#> lm_robust(formula = percent_approved ~ land_ty * decade + land_ceeu *

```

```

#>     decade + land_asia * decade + land_neu_poor * decade + land_other *  

#>     decade + age40 + age60 + age60plus + male + kids + educ_middle +  

#>     educ_high + married + skill_hi + skill_middle + born_ch,  

#>     data = data, se_type = "stata")  

#>  

#> Standard error type: HC1  

#>  

#> Coefficients:  

#>              Estimate Std. Error   t value  

#> (Intercept)    98.4948    5.729  17.19091  

#> land_ty        0.9477    4.115   0.23034  

#> decade1980    -7.4463    4.963  -1.50032  

#> decade1990    -6.5857    4.183  -1.57425  

#> decade2000    -4.7774    4.082  -1.17030  

#> land_ceeu      2.6995    3.675   0.73459  

#> land_asia     -1.0785    5.369  -0.20088  

#> land_neu_poor -102.3791   4.918 -20.81903  

#> land_other     -6.7299    4.389  -1.53337  

#> age40          -4.9521    4.049  -1.22292  

#> age60          -3.5319    4.681  -0.75459  

#> age60plus      -3.0633    6.704  -0.45697  

#> male            1.0254    2.829   0.36252  

#> kids             -4.8851    3.522  -1.38685  

#> educ_middle     6.6233    2.952   2.24342  

#> educ_high       4.8188    4.686   1.02841  

#> married         -0.2325    3.624  -0.06415  

#> skill_hi        1.5583    2.873   0.54233  

#> skill_middle    4.9735    3.660   1.35885  

#> born_ch          0.4051    3.226   0.12559  

#> land_ty:decade1980 -9.3463   8.359  -1.11817  

#> land_ty:decade1990 -47.2903   5.993  -7.89047  

#> land_ty:decade2000 -51.9192   5.684  -9.13391  

#> decade1980:land_ceeu -25.2508  10.050  -2.51260  

#> decade1990:land_ceeu -28.0100  8.807  -3.18055  

#> decade2000:land_ceeu -17.4891  11.888  -1.47113  

#> decade1980:land_asia  11.0299  6.205   1.77754  

#> decade1990:land_asia -2.4578  7.021  -0.35009

```



```

#> decade1990:land_neu_poor    98.9653  121.532 1362
#> decade2000:land_neu_poor    32.0962   93.404 1362
#> decade1980:land_other       0.1040   23.291 1362
#> decade1990:land_other      -1.8825   20.469 1362
#> decade2000:land_other      -0.2687   19.464 1362
#>
#> Multiple R-squared:  0.2947 ,   Adjusted R-squared:  0.2771
#> F-statistic:     NA on 34 and 1362 DF,  p-value: NA

time_fe_results = tidy(mod12)
time_fe_fig_data = data.frame()
for(country in c("ty","ceeu","other","asia","poor")){
  decades=c(1970,1980,1990,2000)
  coeffs = time_fe_results$estimate[grep(country,time_fe_results$term)]
  time_fe_fig_data = bind_rows(time_fe_fig_data,
                                data.frame(decade=decades,
                                           coef=coeffs,
                                           origin=country))
}

time_fe_fig_data = time_fe_fig_data %>%
  mutate(origin=case_when(
    origin=="ty"~"Turkey",
    origin=="ceeu"~"Yugoslavia",
    origin=="other"~"Other",
    origin=="asia"~"Asia",
    origin=="poor"~"NonEuropean Poor"
  )) %>%
  rowwise() %>%
  mutate(num_migrants=100*nrow(data[data$nationality==origin&data$decade==decade,])/nrow(data[data$decade==decade,]))
  ungroup()

time_fe_results = tidy(mod9)

```

```
coeffs = time_fe_results$estimate[20:30]
bar_data = data.frame(term=gsub("land_ty:", "", time_fe_results$term[20:30]), coeff=time_fe_results$estimate[20:30], pval

fig2 = ggplot(bar_data, aes(x=term, y=coeff, fill=(pval<=0.05)))+
  geom_bar(stat="identity") +
  scale_fill_manual(labels=c(FALSE, TRUE), values=c("red", "blue")) +
  ggthemes::theme_fivethirtyeight() +
  theme(axis.title = element_text(), axis.title.x = element_blank()) +
  labs(y="Difference in approval probability relative to all other\nimmigrants (percentage points)")
```

```
ggsave(paste0(charts_folder, "figure2.png"), fig2, width=10, height=5, units="in", dpi=320)
```

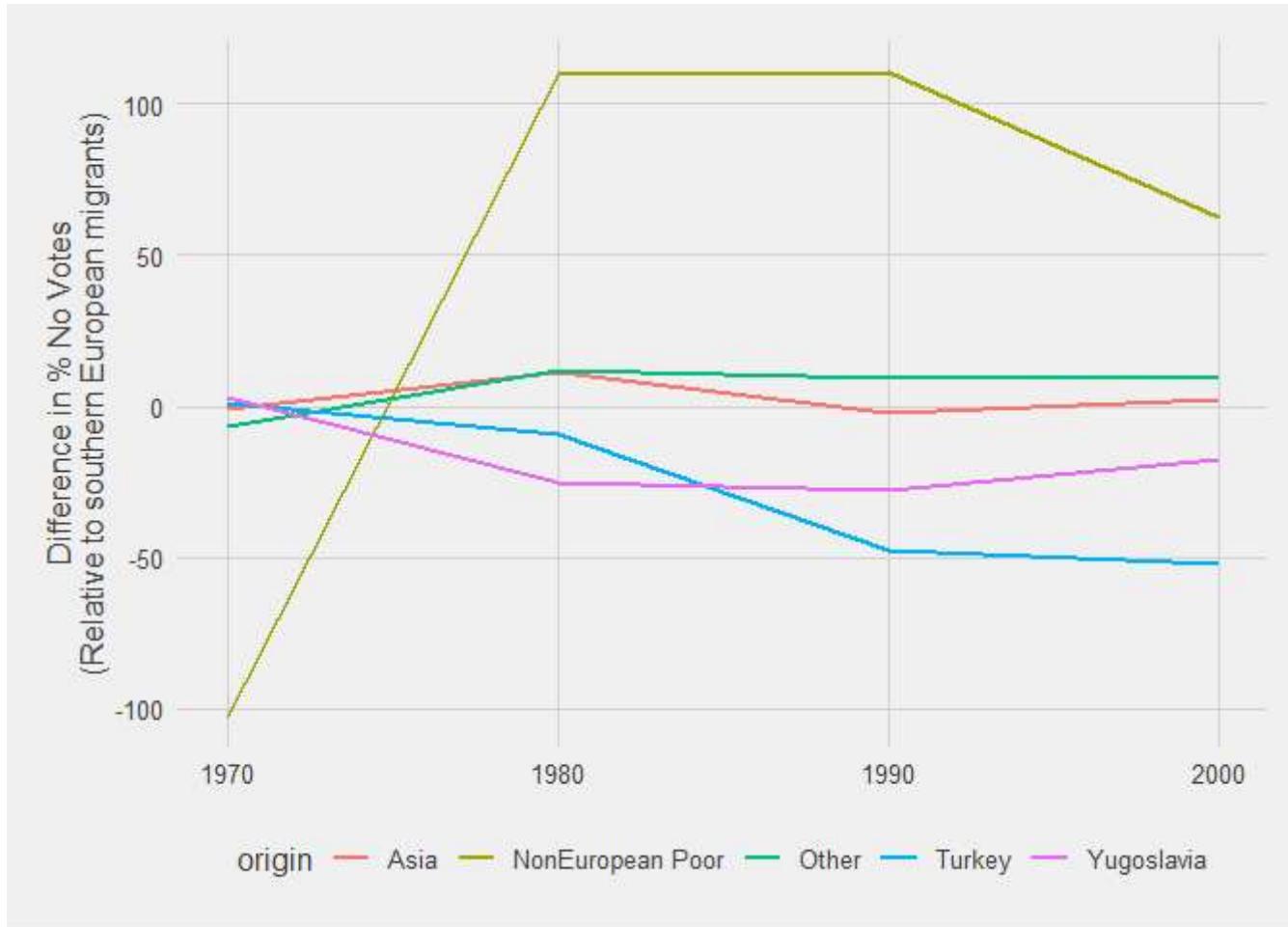
```
time_fe_results = tidy(mod9)
bar_data = data.frame(term=gsub("land_ty:", "", time_fe_results$term[17:19]), coeff=time_fe_results$estimate[17:19], pval

fig3 = ggplot(bar_data, aes(x=term, y=coeff, fill=(pval<=0.05)))+
  geom_bar(stat="identity") +
  scale_fill_manual(labels=c(FALSE, TRUE), values=c("red", "blue")) +
  ggthemes::theme_fivethirtyeight() +
  theme(axis.title = element_text(), axis.title.x = element_blank()) +
  labs(y="Difference in approval probability relative to all other\nimmigrants (percentage points)")
```

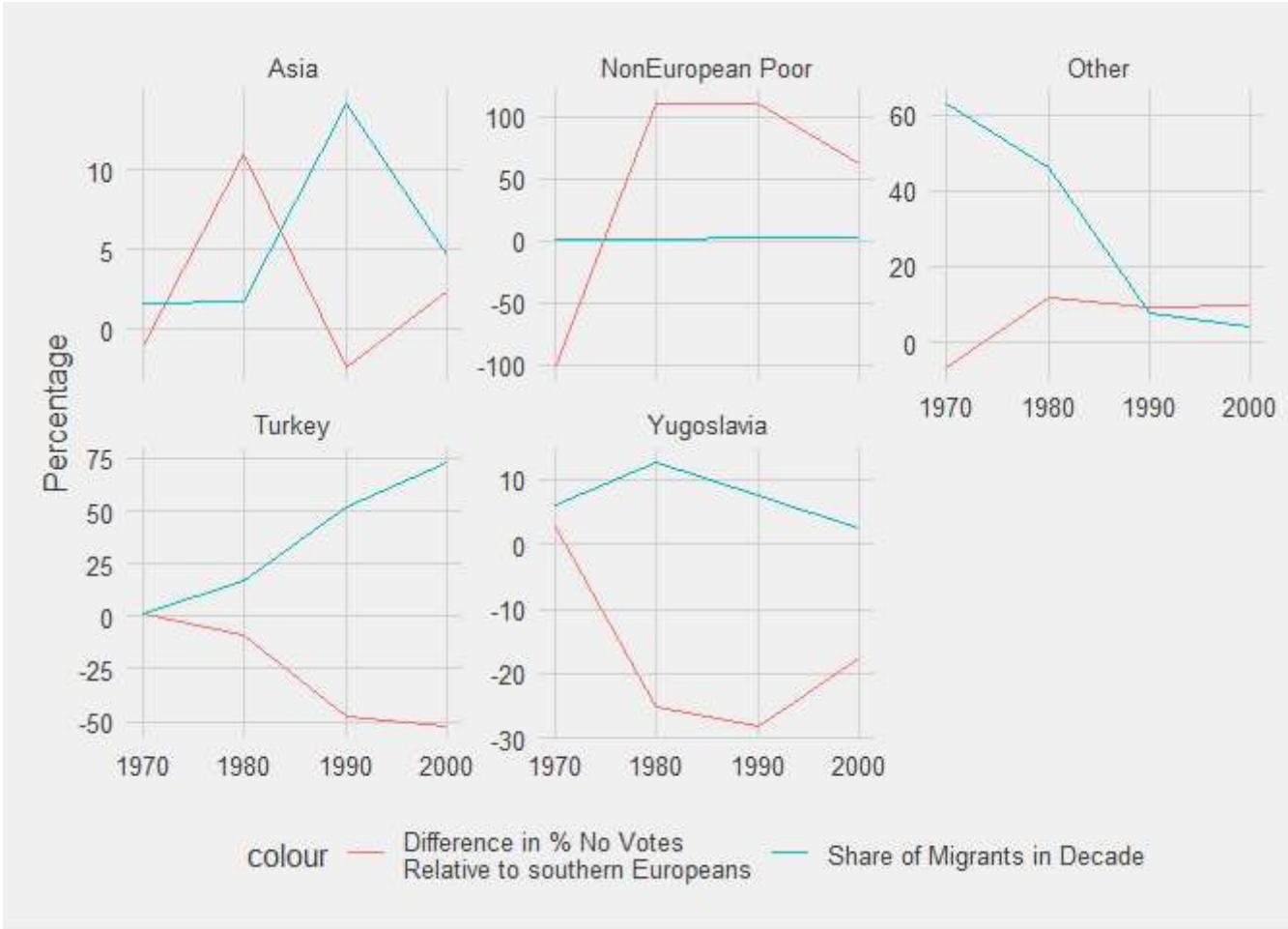
```
ggsave(paste0(charts_folder, "figure3.png"), fig3, width=10, height=5, units="in", dpi=320)
```

```
ggplot(time_fe_fig_data, aes(x=decade, y=coef, color=origin)) +
  geom_line(size=1) +
  labs(x="Decade", y="Difference in % No Votes\n(Relative to southern European migrants)") +
  ggthemes::theme_fivethirtyeight() +
  theme(axis.title = element_text(), axis.title.x = element_blank())
```





```
ggplot(time_fe_fig_data,aes(x=decade)) +
  facet_wrap(~origin,scales="free_y") +
  geom_line(aes(y=coef,color="Difference in % No Votes\nRelative to southern Europeans"))+
  geom_line(aes(y=num_migrants,color="Share of Migrants in Decade")) +
  labs(y="Percentage") +
  ggthemes::theme_fivethirtyeight() +
  theme(axis.title = element_text(), axis.title.x = element_blank())
```



```
# log close
```

Created on 2022-11-30 with [reprex v2.0.2](#)