Appendix: code notebook

Lijphart finds that consensus democracy (on the executive-parties dimension) decreases social inequality. Here are the dependent variables:

1. Gini coefficient;
2. 10/10 ratio;
3. 20/20 ratio;
4. Percentage of women in government;
5. UN Violence against Women statistic.

In this notebook, I aim to do two things:

1. Replicate Lijphart's results using panel data to remove fixed effects that confound the results;
2. Extend Lijphart's research to 67 rather than his original 36 democracies, to see if his results continue to hold.

```
#install.packages(c("psych", "zoo", "plm", "stargazer"))
library("psych")
library("zoo")
```

```
##
## Attaching package: 'zoo'
```

```
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
```

```
library("plm")
```

```
## Loading required package: Formula
```

```
library("stargazer")
```

```
##
## Please cite as:
```

```
##  Hlavac, Marek (2018). stargazer: Well-Formatted Regression and Summary Statistics Tables.
```

```
##  R package version 5.2.1. https://CRAN.R-project.org/package=stargazer
```

```
# Import QoG dataset
qog <- read.csv("./qog.csv") #extremely large
data <- read.csv("http://andy.egge.rs/data/L.csv")
```

Import the Polity IV dataset and take countries that have i) a Democracy score of 6 or more and ii) have been consistently democratic for at least 20 years (1997-2016).

```
# Import Polity IV dataset
polity_iv <- read.csv("./polity_iv_2016.csv")
polity_iv <- polity_iv[, c("scode", "country", "year", "flag", "democ", "autoc", "polity2")]
# These are countries that have been undemocratic at any time since 1996 (Democracy score of <6)

not_democratic = subset(polity_iv, (year>1996 & democ <6))

# The democracies that pass our criteria (current less not democratic)

polity_iv_democracies <- subset(polity_iv, !(is.element(country, not_democratic$country)) & year>1996)
```

```r
# Remove outliers (too young: East Timor, Kosovo, Macedonia, Serbia and Montenegro)
polity_iv_democracies <- subset(polity_iv_democracies, !(is.element(country, "East Timor") |
                                                         is.element(country,"Kosovo") |
                                                         is.element(country, "Serbia") |
                                                         is.element(country, "Montenegro") |
                                                         is.element(country, "Serbia and Montenegro") |
                                                         is.element(country, "Timor Leste")
                                                       ))
# I was debating whether to use Polity score >= 6 or democracy score >= 6. It turns out that
# only Suriname and Albania have a democracy score of 6 with a Polity score of 5 and so
# this doesn't affect the results much
#polity_iv_democracies$country[polity_iv_democracies$polity2<6];

#Cleanup
polity_iv_democracies$country <- factor(polity_iv_democracies$country)
polity_iv_democracies$scode <- factor(polity_iv_democracies$scode)

# Total of 65 countries that fulfill the criteria of at least 20 consecutive
# years of democracy after 1996
polity_iv_democracies_2016 <- subset(polity_iv_democracies, year==2016)
polity_iv_democracies_2016
```

```
##       scode             country year flag democ autoc polity2
## 320     ALB             Albania 2016    0     9     0       9
## 609     ARG           Argentina 2016    1     9     0       9
## 751     AUL           Australia 2016    0    10     0      10
## 968     AUS             Austria 2016    0    10     0      10
## 1352    BEL             Belgium 2016    0     8     0       8
## 1409    BEN               Benin 2016    0     7     0       7
## 1839    BOL             Bolivia 2016    0     7     0       7
## 1915    BOT            Botswana 2016    0     8     0       8
## 2108    BRA              Brazil 2016    0     8     0       8
## 2301    BUL            Bulgaria 2016    0     9     0       9
## 2515    CAN              Canada 2016    0    10     0      10
## 2614    CAP          Cape Verde 2016    0    10     0      10
## 2927    CHL               Chile 2016    0    10     0      10
## 3329    COL            Colombia 2016    0     7     0       7
## 3607    COS          Costa Rica 2016    0    10     0      10
## 3805    CYP              Cyprus 2016    0    10     0      10
## 3904    CZR      Czech Republic 2016    0     9     0       9
## 4121    DEN             Denmark 2016    0    10     0      10
## 4334    DOM  Dominican Republic 2016    1     8     1       7
## 4762    EST             Estonia 2016    0     9     0       9
## 5040    FIN             Finland 2016    0    10     0      10
## 5304    FRN              France 2016    0     9     0       9
## 5747    GMY             Germany 2016    0    10     0      10
## 5980    GRC              Greece 2016    0    10     0      10
## 6184    GUA           Guatemala 2016    1     9     1       8
## 6294    GUY              Guyana 2016    1     8     1       7
## 6669    HON            Honduras 2016    0     7     0       7
## 6819    HUN             Hungary 2016    0    10     0      10
## 6886    IND               India 2016    0     9     0       9
## 7054    IRE             Ireland 2016    0    10     0      10
## 7433    ISR              Israel 2016    0     7     1       6
```

```
## 7589    ITA               Italy 2016    0    10     0      10
## 7704    JAM             Jamaica 2016    0     9     0       9
## 7992    JPN               Japan 2016    0    10     0      10
## 8382    LAT              Latvia 2016    0     8     0       8
## 8792    LIT           Lithuania 2016    0    10     0      10
## 8942    LUX          Luxembourg 2016    0    10     0      10
## 9025    MAC           Macedonia 2016    0     9     0       9
## 9191    MAS           Mauritius 2016    0    10     0      10
## 9439    MEX              Mexico 2016    0     8     0       8
## 9465    MLD             Moldova 2016    0     9     0       9
## 9672    MON            Mongolia 2016    0    10     0      10
## 9984    NAM             Namibia 2016    0     6     0       6
## 10361   NEW         New Zealand 2016    0    10     0      10
## 10540   NIC           Nicaragua 2016    1     7     1       6
## 10857   NOR              Norway 2016    0    10     0      10
## 11059   NTH         Netherlands 2016    0    10     0      10
## 11484   PAN              Panama 2016    0     9     0       9
## 11746   PAR            Paraguay 2016    0     9     0       9
## 12024   PHI         Philippines 2016    0     8     0       8
## 12236   POL              Poland 2016    0    10     0      10
## 12453   POR            Portugal 2016    0    10     0      10
## 12637   ROK         Korea South 2016    0     8     0       8
## 12795   RUM             Romania 2016    0     9     0       9
## 13127   SAF        South Africa 2016    0     9     0       9
## 13303   SAL          El Salvador 2016    0     8     0       8
## 13855   SLO     Slovak Republic 2016    0    10     0      10
## 13881   SLV             Slovenia 2016    0    10     0      10
## 14194   SPN                Spain 2016    0    10     0      10
## 14367   SUR             Suriname 2016    0     6     1       5
## 14633   SWD               Sweden 2016    0    10     0      10
## 14802   SWZ          Switzerland 2016    0    10     0      10
## 14969   TAW               Taiwan 2016    0    10     0      10
## 15380   TRI Trinidad and Tobago 2016    0    10     0      10
## 16019   UKG      United Kingdom 2016    0    10     0      10
## 16247   URU             Uruguay 2016    0    10     0      10
## 16464   USA       United States 2016    1     8     0       8
```

```
# Here we operationalise both the independent and dependent variables
eff_num_parl_parties = "gol_enep"
minimal_winning_one_party_cabinet = "dpi_gf"
executive_dominance_index = NULL
disproportionality_vars = c(
"dpi_gps1",
"dpi_gps2",
"dpi_gps3",
"dpi_gpvs1",
"dpi_gpvs2",
"dpi_gpvs3",
"dpi_gs",
"dpi_ogpvs",
"dpi_nogps",
"dpi_slop1",
"dpi_slop2",
"dpi_slop3",
```

```
"dpi_vslop1",
"dpi_vslop2",
"dpi_vslop3",
"dpi_vsoop",
"dpi_noops",
"dpi_vsul",
"dpi_numul",
"dpi_seats"
)
interest_group_pluralism_index = NULL

controls = c("undp_hdi", "unna_pop")
econ_inequality_vars = c(
"wdi_gini",
"lis_gini",
"wdi_incsh10h",
"wdi_incsh10l",
"wdi_incsh20h",
"wdi_incsh20l"
)
social_inequality_vars = c(
"wdi_lifexp",
"wdi_lifexpf",
"wdi_lifexpm",
"bl_asy15f",
"bl_asy15m",
"bl_asy15mf",
"bl_lh_15f",
"bl_lh_15m",
"bl_lh_15mf",
"vdem_gender",
"ipu_l_s",
"ipu_l_w",
"ipu_u_s",
"ipu_u_w"
)
qog_reduced <- qog[, c(
"ccode",
"cname",
"year",
econ_inequality_vars,
controls,
eff_num_parl_parties,
social_inequality_vars,
disproportionality_vars,
minimal_winning_one_party_cabinet
)]

colnames(qog_reduced)[colnames(qog_reduced) == "gol_enep"] <- "enep"

# Interpolate columns
qog_reduced$bl_asy15f <-
na.approx(qog_reduced$bl_asy15f, na.rm = FALSE)
```

```
qog_reduced$bl_asy15m <-
na.approx(qog_reduced$bl_asy15m, na.rm = FALSE)
qog_reduced$bl_asy15mf <-
na.approx(qog_reduced$bl_asy15mf, na.rm = FALSE)
qog_reduced$bl_lh_15f <-
na.approx(qog_reduced$bl_lh_15f, na.rm = FALSE)
qog_reduced$bl_lh_15m <-
na.approx(qog_reduced$bl_lh_15m, na.rm = FALSE)
qog_reduced$bl_lh_15mf <-
na.approx(qog_reduced$bl_lh_15mf, na.rm = FALSE)
```

```
# =====
# Here I clean up the data: I rename countries so that they are consistent in
# both datasets and finally merge both datasets together
# =====

colnames(qog_reduced)[colnames(qog_reduced)=="cname"] <- "country"

#Rename Cyprus, France, South Korea and Slovak Republic
levels(qog_reduced$country)[46] <- "Cyprus"
levels(qog_reduced$country)[64] <- "France"
levels(qog_reduced$country)[96] <- "Korea South"
levels(qog_reduced$country)[162] <- "Slovak Republic"
qog_democracies <-
  subset(qog_reduced, (
  is.element(country, polity_iv_democracies_2016$country) &
  year > 1996 & year <= 2016
  ))

# Cleanup
qog_democracies$country <- factor(qog_democracies$country)
qog_democracies_2016 <- subset(qog_democracies, year==2016)

# Sanity check
stopifnot(levels(polity_iv_democracies$country) == levels(qog_democracies$country))

# Merge the two data sets together
democracies <- merge(qog_democracies, polity_iv_democracies, by=c("country", "year"))

democracies$dpi_ogpvs[is.na(democracies$dpi_ogpvs)] <- 0 #vote share of other government parties
democracies$dpi_gpvs3 <-na.locf(democracies$dpi_gpvs3)
democracies$dpi_vslop3 <-na.locf(democracies$dpi_vslop3)
democracies$dpi_vsoop <-na.locf(democracies$dpi_vsoop)
```

After merging and cleaning the dataset, I can finally start operationalising my dependent variables. I calculate Gallagher's index of disproportionality.

```
# Calculate Gallagher's index
democracies$disproportionality <- sqrt(0.5 * (
  (democracies$dpi_gpvs1 - democracies$dpi_gps1/democracies$dpi_seats)^2   +
  (democracies$dpi_gpvs2 - democracies$dpi_gps2/democracies$dpi_seats)^2   +
  (democracies$dpi_gpvs3 - democracies$dpi_gps3/democracies$dpi_seats)^2   +
  (democracies$dpi_ogpvs - democracies$dpi_nogps/democracies$dpi_seats)^2  +
  (democracies$dpi_vslop1 - democracies$dpi_slop1/democracies$dpi_seats)^2 +
```

```
  (democracies$dpi_vslop2 - democracies$dpi_slop2/democracies$dpi_seats)^2 +
  (democracies$dpi_vslop3 - democracies$dpi_slop3/democracies$dpi_seats)^2 +
  (democracies$dpi_vsoop - democracies$dpi_noops/democracies$dpi_seats)^2 +
  (democracies$dpi_vsul - democracies$dpi_numul/democracies$dpi_seats)^2
  ))
```

```
# Create the 10/10 and 20/20 ratio
democracies$s10 <- democracies$wdi_incsh10h / democracies$wdi_incsh10l
democracies$s20 <- democracies$wdi_incsh20h / democracies$wdi_incsh20l
```

```
# Create the percentage of women in upper and lower house
democracies$lw <- democracies$ipu_l_w / democracies$ipu_l_s
democracies$uw <- democracies$ipu_u_w / democracies$ipu_u_s
```

```
# Calculate Cronbach's alpha for my exec-parties index

exec_parties <- democracies[c("dpi_gf", "enep", "disproportionality")]

# Reverse the direction of disproportionality as the
# more disproportinate, the more majoritarian
summary(alpha(scale(exec_parties), keys=c("disproportionality")))
```

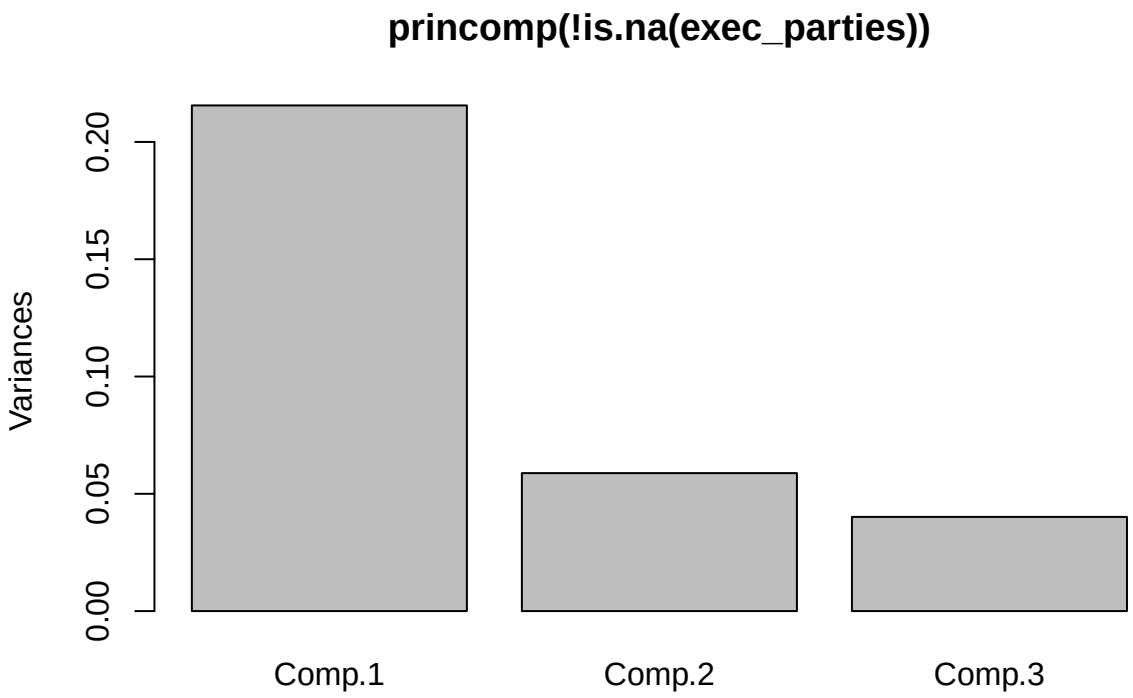
```
##
## Reliability analysis
##  raw_alpha std.alpha G6(smc) average_r S/N    ase mean   sd
##       0.85      0.85    0.79      0.65 5.5 0.0073 0.62 0.98
```

```
# My panel data version of Lijphart's executive-parties dimension is highly
# internally consistent with a Cronbach's alpha of 0.85
```

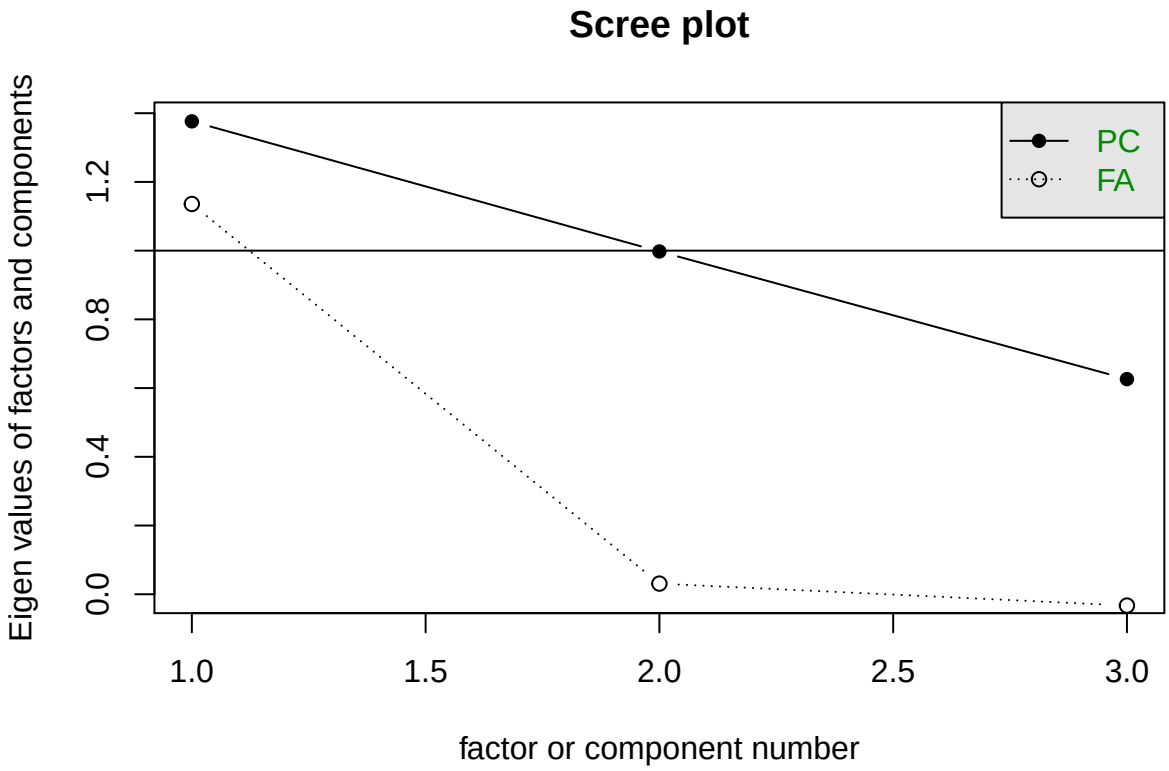
I now run a panel-data regression with my executive-parties dimension (disproportionality, dpi_gf and dpi_enep) as the independent variable and Gini coefficient (wdi_gini) as the dependent variable, controlling for HDI and population size (undp_hdi, unna_pop).

```
# Clean data by scaling everything
d2 <- democracies

d2$wdi_gini <-ifelse(!is.na(d2$wdi_gini), d2$wdi_gini, d2$lis_gini * 100)
d2$undp_hdi <- scale(d2$undp_hdi)
d2$log_unna_pop <- scale(log(d2$unna_pop))

panel <- pdata.frame(d2)
```

```
# Use principal components analysis to reduce to one executive-parties dimension.

# Use a scree plot to see how many factors we need

screeplot(princomp(!is.na(exec_parties)))
```




```
scree(!is.na(exec_parties))
```




```
# The variation explained drops dramatically after 1 component; 1 component is enough.
ep <- principal(exec_parties, nfactors = 1, rotate="none", scores = T)
ep
```

```
## Principal Components Analysis
## Call: principal(r = exec_parties, nfactors = 1, rotate = "none", scores = T)
## Standardized loadings (pattern matrix) based upon correlation matrix
##                       PC1   h2   u2 com
## dpi_gf               0.88 0.77 0.23   1
## enep                 0.86 0.74 0.26   1
## disproportionality -0.91 0.84 0.16   1
##
##                 PC1
## SS loadings    2.35
## Proportion Var 0.78
##
## Mean item complexity =  1
## Test of the hypothesis that 1 component is sufficient.
##
## The root mean square of the residuals (RMSR) is  0.11
##  with the empirical chi square  100.9  with prob <  NA
##
## Fit based upon off diagonal values = 0.97
```

```
#ep$scores
panel$ep <- ep$scores
```

```r
# Create Lijphart's 36 democracies
panel_lijphart <- subset(panel, ccode == 32 | ccode == 36 | ccode == 40 |
ccode == 56 | ccode == 124 | ccode == 188 | ccode == 208 | ccode == 246 |
ccode == 250 | ccode == 276 | ccode == 300 | ccode == 356 | ccode == 372 |
ccode == 376 | ccode == 380 | ccode == 388 | ccode == 392 | ccode == 410 |
ccode == 480 | ccode == 528 | ccode == 554 | ccode == 578 | ccode == 620 |
ccode == 724 | ccode == 752 | ccode == 756 | ccode == 780 | ccode == 826 |
ccode == 840 | ccode == 858)
```

```r
# Test the hypothesis that Lijphart's countries are more gender equal than my countries
panel_non_lijphart <- subset(panel, !(ccode %in% panel_lijphart$ccode))
t.test((panel_lijphart$vdem_gender), (panel_non_lijphart$vdem_gender))
```

```
##
##  Welch Two Sample t-test
##
## data:  (panel_lijphart$vdem_gender) and (panel_non_lijphart$vdem_gender)
## t = 17.951, df = 1317.1, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.07289254 0.09077915
## sample estimates:
## mean of x mean of y
## 0.8976427 0.8158068
```

```r
# Remove Argentina as a inflation outlier
panel_lijphart <- subset(panel_lijphart, ccode!=32)

# Baseline panel regressions without any fixed effect controlling
gini_lijphart_bl <-
  lm(
  panel_lijphart$wdi_gini ~ panel_lijphart$ep + panel_lijphart$undp_hdi + panel_lijphart$log_unna_pop
  )
  s10_lijphart_bl <-
  lm(
  panel_lijphart$s10 ~ panel_lijphart$ep + panel_lijphart$undp_hdi + panel_lijphart$log_unna_pop
  )
  s20_lijphart_bl <-
  lm(
  panel_lijphart$s20 ~ panel_lijphart$ep + panel_lijphart$undp_hdi + panel_lijphart$log_unna_pop
  )
  lw_lijphart_bl <-
  lm(
  panel_lijphart$lw ~ panel_lijphart$ep + panel_lijphart$undp_hdi + panel_lijphart$log_unna_pop
  )
  uw_lijphart_bl <-
  lm(
  panel_lijphart$uw ~ panel_lijphart$ep + panel_lijphart$undp_hdi + panel_lijphart$log_unna_pop
  )
  se_lijphart_bl <-
  lm(
  panel_lijphart$vdem_gender ~ panel_lijphart$ep + panel_lijphart$undp_hdi + panel_lijphart$log_unna_po
  )
```

```r
gini_lijphart <-
plm(
panel_lijphart$wdi_gini ~ panel_lijphart$ep + panel_lijphart$undp_hdi + panel_lijphart$log_unna_pop,
panel_lijphart,
effect = "twoways",
method = "within"
)
s10_lijphart   <-
plm(
panel_lijphart$s10 ~ panel_lijphart$ep + panel_lijphart$undp_hdi + panel_lijphart$log_unna_pop,
panel_lijphart,
effect = "twoways",
method = "within"
)
s20_lijphart   <-
plm(
panel_lijphart$s20 ~ panel_lijphart$ep + panel_lijphart$undp_hdi + panel_lijphart$log_unna_pop,
panel_lijphart,
effect = "twoways",
method = "within"
)
lw_lijphart    <-
plm(
panel_lijphart$lw ~ panel_lijphart$ep + panel_lijphart$undp_hdi + panel_lijphart$log_unna_pop,
panel_lijphart,
effect = "twoways",
method = "within"
)
uw_lijphart    <-
plm(
panel_lijphart$uw ~ panel_lijphart$ep + panel_lijphart$undp_hdi + panel_lijphart$log_unna_pop,
panel_lijphart,
effect = "twoways",
method = "within"
)
se_lijphart    <-
plm(
panel_lijphart$vdem_gender ~ panel_lijphart$ep + panel_lijphart$undp_hdi + panel_lijphart$log_unna_pop
panel_lijphart,
effect = "twoways",
method = "within"
)


# stargazer(
#   gini_lijphart_bl,
#   s10_lijphart_bl,
#   s20_lijphart_bl,
#   #report="vc*p",
#   title = "Baseline panel regression for Lijphart's 36 countries: Economic inequality",
#   dep.var.labels = c("Gini", "10/10 ratio", "20/20 ratio"),
#   covariate.labels = c("Executive-parties", "HDI", "Population, logged"),
#   type = "latex"
```

```
  #   )
  #   stargazer(
  #   lw_lijphart_bl,
  #   uw_lijphart_bl,
  #   #report="vc*p",
  #   title = "Baseline panel regression for Lijphart's 36 countries: Gender inequality",
  #   dep.var.labels = c("Women in lower house", "Women in upper house", "Gender inequality"),
  #   covariate.labels = c("Executive-parties", "HDI", "Population, logged"),
  #   type = "latex"
  #   )
  #   stargazer(
  #   lw_lijphart,
  #   uw_lijphart,
  #   #report="vc*p",
  #   title = "Fixed effect panel regression for Lijphart's 36 countries: Gender inequality",
  #   dep.var.labels = c("Women in lower house", "Women in upper house", "Gender inequality"),
  #   covariate.labels = c("Executive-parties", "HDI", "Population, logged"),
  #   type = "latex"
  #   )
  #
  #   stargazer(
  #   gini_lijphart,
  #   s10_lijphart,
  #   s20_lijphart,
  #   title = "Fixed effect panel regression for Lijphart's 36 countries: Economic inequality",
  #   dep.var.labels = c("Gini", "10/10 ratio", "20/20 ratio"),
  #   covariate.labels = c("Executive-parties", "HDI", "Population, logged"),
  #   type = "text"
  #   )
```

```
gini  <-
  plm(
  panel$wdi_gini ~ panel$ep + panel$undp_hdi + panel$log_unna_pop,
  panel,
  effect = "twoways",
  method = "within"
  )
  s10   <-
  plm(
  panel$s10 ~ panel$ep + panel$undp_hdi + panel$log_unna_pop,
  panel,
  effect = "twoways",
  method = "within"
  )
  s20   <-
  plm(
  panel$s20 ~ panel$ep + panel$undp_hdi + panel$log_unna_pop,
  panel,
  effect = "twoways",
  method = "within"
  )


  lw    <-
```

```r
plm(
panel$lw ~ panel$ep + panel$undp_hdi + panel$log_unna_pop
,
panel,
effect = "twoways",
method = "within"
)
uw    <-
plm(
panel$uw ~ panel$ep + panel$undp_hdi + panel$log_unna_pop,
panel,
effect = "twoways",
method = "within"
)
se    <-
plm(
panel$vdem_gender ~ panel$ep + panel$undp_hdi + panel$log_unna_pop,
panel,
effect = "twoways",
method = "within"
)

baseline_lw <-
lm(panel$lw ~ panel$ep + panel$undp_hdi + panel$log_unna_pop)
baseline_uw <-
lm(panel$uw ~ panel$ep + panel$undp_hdi + panel$log_unna_pop)
baseline_se <-
lm(panel$vdem_gender ~ panel$ep + panel$undp_hdi + panel$log_unna_pop)

# stargazer(
#   gini,
#   s10,
#   s20,
#   title = "Fixed effect panel regression for extended 63 countries: Economic inequality",
#   dep.var.labels = c("Gini", "10/10 ratio", "20/20 ratio"),
#   covariate.labels = c("Executive-parties", "HDI", "Population, logged"),
#   type = "text"
#   )
#
#   stargazer(
#   baseline_lw,
#   baseline_uw,
#   type = "latex",
#   title = "Baseline panel regression for extended 63 countries: Gender inequality",
#   dep.var.labels = c("Women in lower house", "Women in upper house", "Gender inequality"),
#   covariate.labels = c("Executive-parties", "HDI", "Population, logged")
#   )
#
#   stargazer(
#   lw,
#   uw,
#   se,
#   type = "text",
```

```
#   title = "Fixed effect panel regression for extended 63 countries: Gender inequality",
#   dep.var.labels = c("Women in lower house", "Women in upper house", "Gender inequality"),
#   covariate.labels = c("Executive-parties", "HDI", "Population, logged")
#   )
```