Does consensus democracy reduce social inequality?

Zhenghong Lieu

1. Introduction

Lijphart (2012) argues that consensus democracy reduces social inequality due to increased representation of minority voices, but Tsebelis (2002) provides that the profusion of political actors could also cause policy gridlock[1] and stymie the decisive political action needed for redistribution.

Prima facie, Lijphart's regressions seem to provide strong empirical evidence for his theory. However in this essay, I demonstrate that Lijphart's regression results cannot be taken at face value due to the presence of time- and country- based heterogeneities as well as the methodological flaw of selecting on the dependent variable.

I then use fixed-effect panel regressions that control for unobserved heterogeneities at the time- and country- level to attempt to replicate and extend Lijphart's findings more robustly. I begin by replicating Lijphart's analysis on his original 36 democracies, then expand the analysis to include 31 new democracies that have since fulfilled Lijphart's criteria of 20 years of continuous democracy.

Due to the lack of data I could not replicate Lijphart's results on economic inequality. However, I successfully replicate Lijphart's results on gender inequality at the 1% significance level—and show that after controlling for heterogeneity and/or extending his analysis to include new democracies, these results are no longer significant.

I am unable to make claims about consensus democracy *vis-a-vis* economic inequality but can definitively reject the claim that consensus democracy reduces gender inequality. I also conclude that there is an urgent need for free and open data in the interests of conducting and replicating key results in political science.

2. Theoretical argument

Lijphart (2012) argued that the executive-parties dimension of consensus democracies (greater effective number of parties, less government fractionalisation and electoral disproportionality) better incorporate multiple voices and points of

[1] Conditional on moderate to high levels of ideological polarisation and internal policy congruence of veto players.

view. As such, consensus democracies can better serve the needs of not just the majority but also underprivileged and marginalised minorities, and in this way reduce social inequality.

On the other hand, it could be the case that consensus democracies are less decisive and cannot enact the large and substantial political change needed for true social redistribution. Tsebelis (2002) provides a theoretical basis for such a claim in *Veto Players*: the more *veto players* (political actors whose agreement is necessary to enact political change) are present in a system, the less likely it is that there will be an acceptable policy outcome, and policy gridlock will result.

Tsebelis's veto players framework is well-supported by empirical evidence: Boix et al. (2007) found that policy change is "...slower and less dramatic under presidential systems than parliamentary ones, all else equal" due to the addition of an extra veto player (the president). Consensual democracies exhibit greater legislative heterogeneity, as they have electoral systems that allow smaller, fringe political parties to win seats in the legislature. Legislative heterogeneity has been shown to lead to policy gridlock (see Binder, 1999). For these reasons, consensus democracies may fail to reduce social inequality if policy gridlock prevents governments from enacting the broad, sweeping redistributive policies that reduce social inequality.

Both theories are plausible, so we must look to the empirical evidence to distinguish between the two. I now look at Lijphart's findings and explain why, due to several methodological flaws, they cannot be accepted as evidence for the claim.

3. Methodological criticisms

Lijphart runs many multivariate regressions and finds that consensus democracy (operationalised by the executive-parties dimension) is negatively correlated with social inequality. However there are three confounders that reduce the validity of his data.

First of all, Lijphart's data suffers from country-level heterogeneity. While social inequality can be mediated through political systems, it is to a large extent dependent on many historical and cultural factors that may drive both the formation of consensual democratic institutions and policies that reduce social inequality. Suppose that a group of countries possess some common factor (climate, culture, public *zeitgeist*) that causes them to pursue consensual institutions and reduce social inequality; for example, Japan's communitarian culture results in both consensual political behaviour and redistributive social norms. Then a spurious correlation would emerge between consensus democracy and social inequality. Introducing fixed effects removes the influence of any unknown common factor.

Time-level heterogeneity is also present. For example, taking the Gini coefficient circa 2000 leaves him open to criticism regarding the choice of time period. Lijphart used the Gini coefficient circa 2000, the peak of the dot-com boom. It may be the case that his relationship no longer hold in a recession, when only the

decisiveness of majoritarian systems can enact stimulus spending.[2] To rule out such a possibility, therefore, panel data must be used to control for time-level heterogeneities.

Lastly, Lijphart's case selection falls prey to *selecting on the dependent variable*. In his case selection, Lijphart selects only countries that have a Freedom House rating of "Free". This is problematic as Freedom House is a substantive measure that includes civil liberties in its operationalisation of democracy. That Civil Liberties index includes the requirement that a state must strive for "equality of opportunity for everyone, including women and minority groups"[3]. By choosing only "Free" countries, Lijphart is guilty of selecting on the dependent variable: choosing only countries that have a certain degree of social equality already.

I have identified confounders that reduce the validity of Lijphart's regressions. Fortunately, conducting panel analysis with unobserved fixed effects can ameliorate the effects of these confounders. However, using panel data requires the construction of new time-series variables—the details of which I elucidate now.

4. Operationalisation

To conduct panel data analysis, I must first construct time-series analogues of the original variables. Specifically, I analyse 20 years of data (1997–2016) for 36 and 67 countries.

I use the Quality of Government 2018 (QoG) Standard Dataset, which contains "approximately 2100 variables from more than 100 data sources" (Teorell, 2018), to create my own executive-parties dimension that hews closely to Lijphart's. To replicate Lijphart's results as closely as possible, and to prove that my panel data is highly relevant to Lijphart's regressions, I use all but one of Lijphart's dependent variables.

Executive-parties dimension

I follow William Lai's (Lai, 2017) example and operationalise the executive-parties dimension in the same way as he does. Specifically, I exclude the measures of executive dominance and interest-group pluralism, and substitute government fractionalisation for percentage minimal-winning cabinets.

Effective number of parties

I take this variable from the QoG dataset.

Electoral disproportionality

I pull raw data from the QoG dataset to calculate Gallagher's index of disproportionality myself.

[2] As mentioned earlier, consensus democracies may fail to enact stimulus spending if they tend towards policy gridlock.

[3] "Civil Liberties" section of https://freedomhouse.org/report/methodology-freedom-world-2018

Percentage minimal-winning cabinets

As the QoG dataset does not contain data about minimal-winning cabinets, I use the index of government fractionalisation instead. "Dodd (1976, p. 133), whose definitions Lijphart (2012) uses, has shown government fractionalization to be a good proxy for power sharing and deviation from minimal-winning cabinet status." (Lai, 2017).

Executive dominance

I excluded the measure of executive dominance because I agree with Lai's criticism of Lijphart's operationalisation of executive dominance—Lijphart uses his intuition to give arbitrary, impressionistic values to non-parliamentary governments.

Interest-group pluralism

I excluded interest-group pluralism as it exhibits completely different correlational behaviour compared to the other four variables (see Giuliani, 2016).

Running confirmatory principal components analysis (PCA) on the variables shows that the variation explained drops dramatically after one factor [1], showing that these metrics do indeed reduce to a single executive-parties dimension. This dimension also has high internal consistency with a Cronbach's α of 0.85, which further bolsters its validity.

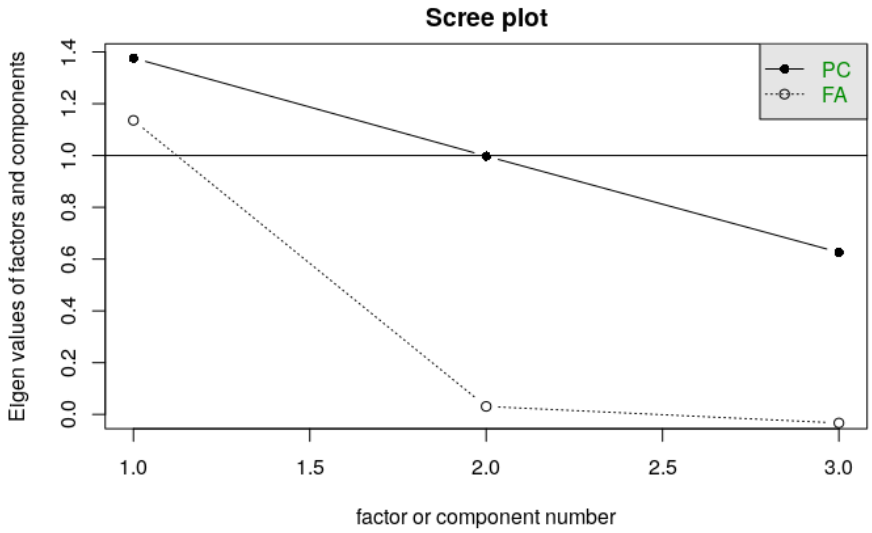


Figure 1: Scree plot of executive-parties dimension

Social inequality

I mostly follow Lijphart's dependent variables for social inequality: Gini coefficient, 10/10 and 20/20 ratio and women in upper/lower houses. I do not use

the UN Gender Inequality Index, as it is only measured every five years and is unavailable in the QoG dataset.

Extending Lijphart's analysis to 67 democracies

I also extend the sample of democracies to include new democracies that, as of 2016, fulfill Lijphart's original criteria for country selection:

```
1. Population greater than 500,000;
2. At least 20 years of continuous democracy.
```

I did this because Lijphart's originally selected only those countries that were "Free" on the Freedom House index, which makes him guilty of *selecting on the dependent variable*: choosing only countries that have a certain degree of social equality already. Therefore I avoid this by using the Polity IV data series instead and select countries with a score of 6 or more for the Democracy index, which gives us 67 democracies. For these 67 countries, I take 20 years of time-series data from 1997 to 2016, which gives us a total of N=1340 observations.

5. Methodology

I run panel data regressions on the following dependent variables, both on Lijphart's original 36 and my extended 67 democracies:

```
1. Gini coefficient;
2. 10/10 ratio;
3. 20/20 ratio;
4. Percentage of women in upper house;
5. Percentage of women in lower house.
```

Lijphart's 36 countries

I first run baseline panel regressions on Lijphart's 36 countries (N=580), without controlling for fixed effects. I fail to replicate Lijphart's findings on Gini, 10/10 and 20/20 ratio (economic inequality). Table 1 details these findings. There is a relationship between consensus democracy and Gini at the 5% level but no relationships for 10/10 and 20/20 ratios, so results are mixed at best.

On the other hand, I replicate Lijphart's findings on gender inequality perfectly. The relationships between consensus democracy and gender inequality are all highly significant at the 1% level and in some cases even the 0.05% level.

My successful replication of Lijphart's findings suggest that my constructed executive-parties and gender inequality variables hew closely to Lijphart's and are thus highly applicable in engaging with his results.

I then start controlling for time- and country-level heterogeneity (`effect="twoways", method="within"`), as well as HDI and logged population. As Lijphart correctly identified, HDI and population can confound the results. While my model can

Table 1: Baseline panel regression for Lijphart's 36 countries: Economic inequality

	Dependent variable:		
	Gini	10/10 ratio	20/20 ratio
	(1)	(2)	(3)
Executive-parties	-0.736^{**}	0.320	0.114
	(0.353)	(0.366)	(0.147)
HDI	-10.700^{***}	-12.241^{***}	-5.111^{***}
	(0.664)	(0.674)	(0.271)
Population, logged	0.617	0.576	0.256
	(0.465)	(0.487)	(0.196)
Constant	42.125^{***}	21.100^{***}	10.431^{***}
	(0.546)	(0.560)	(0.225)
Observations	219	197	197
R^2	0.668	0.716	0.731
Adjusted R^2	0.663	0.711	0.727
Residual Std. Error	3.897 (df = 215)	3.799 (df = 193)	1.529 (df = 193)
F Statistic	144.241^{***} (df = 3; 215)	161.977^{***} (df = 3; 193)	175.086^{***} (df = 3; 193)
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 2: Baseline panel regression for Lijphart's 36 countries: Gender inequality

	Dependent variable:	
	Women in lower house	Women in upper house
	(1)	(2)
Executive-parties	0.028^{***}	0.018^{***}
	(0.005)	(0.005)
HDI	0.067^{***}	0.040^{***}
	(0.009)	(0.009)
Population, logged	-0.037^{***}	-0.034^{***}
	(0.005)	(0.005)
Constant	0.205^{***}	0.216^{***}
	(0.008)	(0.008)
Observations	440	267
R^2	0.303	0.226
Adjusted R^2	0.298	0.217
Residual Std. Error	0.095 (df = 436)	0.084 (df = 263)
F Statistic	63.252^{***} (df = 3; 436)	25.535^{***} (df = 3; 263)
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 3: Fixed effect panel regression for Lijphart's 36 countries: Gender inequality

	Dependent variable:	
	Women in lower house	Women in upper house
	(1)	(2)
Executive-parties	0.001	0.015
	(0.005)	(0.012)
HDI	-0.041^{*}	-0.196^{***}
	(0.024)	(0.047)
Population, logged	0.089	0.570^{***}
	(0.091)	(0.205)
Observations	440	267
R^2	0.010	0.091
Adjusted R^2	−0.112	−0.060
F Statistic	1.264 (df = 3; 391)	7.636^{***} (df = 3; 228)
Note:		*p<0.1; **p<0.05; ***p<0.01

control for HDI and population, I include them to allow the coefficients to more accurately represent the true country- and time-heterogeneity. After doing so, I find that the effects of consensus democracy on gender inequality become statistically insignificant.

My extended 67 countries

I ran baseline panel regressions on the 67 countries and find no statistically significant relationships whatsoever. Table [4] details the results of these regressions. (regressions on economic inequality not shown). As we can see, Lijphart's results do not extend to democracies other than his chosen 36.

Table 4: Baseline panel regression for extended 67 countries: Gender inequality

	Dependent variable:	
	Women in lower house	Women in upper house
	(1)	(2)
Executive-parties	0.004	0.006
	(0.003)	(0.004)
HDI	0.060***	0.062***
	(0.004)	(0.006)
Population, logged	−0.016***	−0.023***
	(0.003)	(0.005)
Constant	0.188***	0.172***
	(0.003)	(0.005)
Observations	846	400
R^2	0.253	0.268
Adjusted R^2	0.250	0.262
Residual Std. Error	0.090 (df = 842)	0.087 (df = 396)
F Statistic	95.082*** (df = 3; 842)	48.215*** (df = 3; 396)
Note:		*p<0.1; **p<0.05; ***p<0.01

7. Conclusion

Gender inequality

My regressions point toward a strong refutation of Lijphart's results for the gender inequality dimension. Consensus democracy improves neither women's representation nor political empowerment once I:

i) include a larger sample of 67 countries which have not been preselected for gender equality and/or
ii) control for country- and time-level heterogeneities.

Most interestingly, the relationship between consensus democracy and gender inequality disappears once we include a larger sample of 67 countries (see Table [4]). This lends strong credence to my initial hypothesis that Lijphart was selecting on the dependent variable when he chose countries that were "Free" on the Freedom House metric. My data show exactly this: women in Lijphart's 36 democracies enjoy greater fundamental civil liberties and political empowerment (0.899) compared to women in the 27 (0.817).[4]

Due to this fact, if any relationship continues to hold, then it must be the more nuanced claim that consensus democracy is correlated with gender equality *after a certain level of gender equality is achieved*. But in this case the causal arrow may be reversed! Suppose that some factor of women's biology or socialisation causes women to pursue consensual, prosocial measures in the political arena (Soutschek et al., 2017). Then greater gender equality (more women in parliament) directly causes consensus democracy, but this is a threshold effect only observed once women and men achieve a certain level of gender equality necessary for women to run for—and be in—office. Testing this reverse causal relationship could be done with instrumental variable estimation[5] (see Cederman et al., 2015) but this is beyond the scope of this essay.

Economic inequality

I was unable to replicate Lijphart's results for economic inequality because of the paucity of the data. For instance, even the best source of Gini coefficient data that I could access—the World Bank—is missing data for two-thirds of the observations, causing the panel dataset to be highly unbalanced.

Worse, this missing data is *systematically* unbalanced in two ways:

i. less economically developed countries are more likely to take loans from the World Bank, and they provide the World Bank with Gini estimates;
ii. nations in the European Union all publish Gini estimates.

This panel data is thus systematically biased towards representing poorer, more profligate nations as well as EU member states, which definitely affect replication. Most damningly, the Gini coefficient data for was missing 13 out of 30 values for the year 2000, which prevented me from replicating even his cross-section regression. The Economist Intelligence Unit (EIU) may have better Gini data (Lijphart used it to get the Gini coefficient circa 2000) but it is locked behind a paywall and costs £800 per year to access.

[4]This difference is highly significant: t-value 17.8, p-value $< 2.2e16$

[5]or regression discontinuity, difference-in-differences models, and experimental research

Final thoughts

My results show conclusively that consensus democracy does not reduce gender inequality, however I cannot reject the claim for economic inequality. My failure to replicate Lijphart's findings for the economic inequality dimension highlights the urgent need for comprehensive, free and open data. Without robust data, many of the key findings in political science may suffer from a crisis of replication.

References

Binder, S.A., 1999. The dynamics of legislative gridlock, 1947–96. American Political Science Review 93, 519–533. https://doi.org/10.2307/2585572

Boix, C., Stokes, S.C., Samuels, D., 2007. Separation of powers.

Cederman, L.-E., Hug, S., Schädel, A., Wucherpfennig, J., 2015. Territorial autonomy in the shadow of conflict: Too little, too late? American Political Science Review 109, 354–370. https://doi.org/10.1017/S0003055415000118

Croissant, Y., Millo, G., 2008. Panel data econometrics in r: The plm package. Journal of Statistical Software, Articles 27, 1–43. https://doi.org/10.18637/jss.v027.i02

Giuliani, M., 2016. Patterns of democracy reconsidered: The ambiguous relationship between corporatism and consensualism. European Journal of Political Research 55, 22–42. https://doi.org/10.1111/1475-6765.12117

Lai, W.N., 2017. Does consensus democracy improve economic outcomes?

Lijphart, A., 2012. Patterns of democracy. Yale University Press.

Soutschek, A., Burke, C.J., Raja Beharelle, A., Schreiber, R., Weber, S.C., Karipidis, I.I., Velden, J. ten, Weber, B., Haker, H., Kalenscher, T., Tobler, P.N., 2017. The dopaminergic reward system underpins gender differences in social preferences. Nature Human Behaviour 1, 819–827. https://doi.org/10.1038/s41562-017-0226-y

Teorell, S.D., Jan, 2018. The quality of government standard dataset, version jan18. https://doi.org/doi:10.18157/QoGStdJan18

Tsebelis, G., 2002. Veto players: How political institutions work. Princeton University Press.

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")))
```

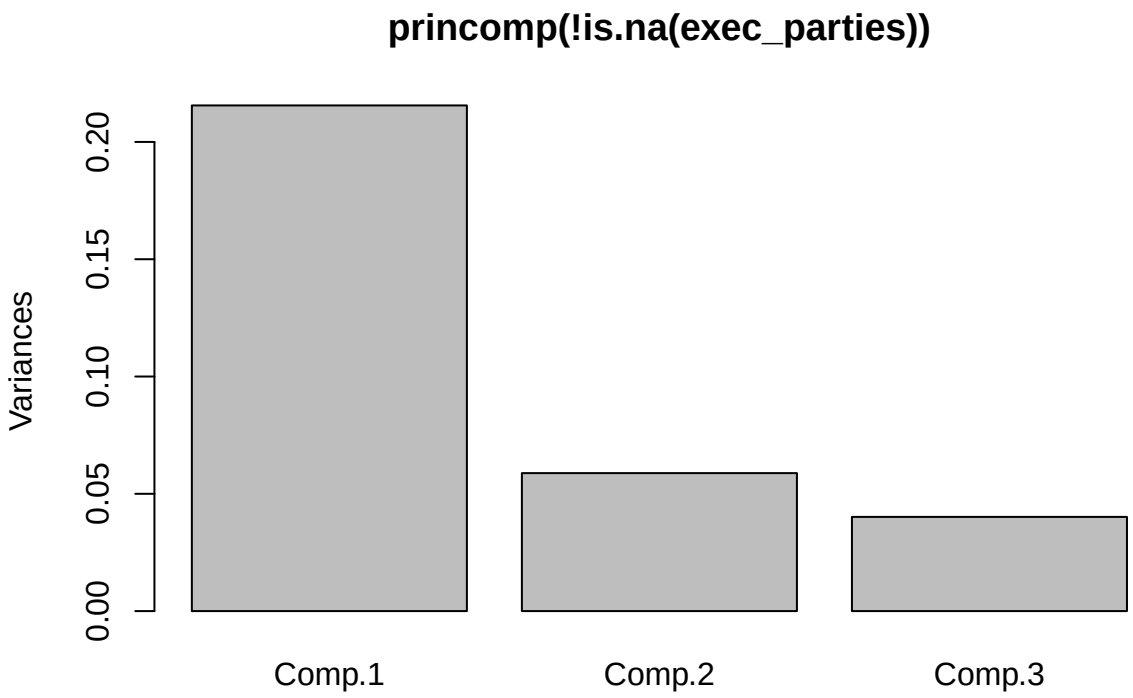
```
##
## Reliability analysis
##  raw_alpha std.alpha G6(smc) average_r S/N    ase mean   sd
##       0.85      0.85    0.79      0.65 5.5 0.0073 0.62 0.98
```

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

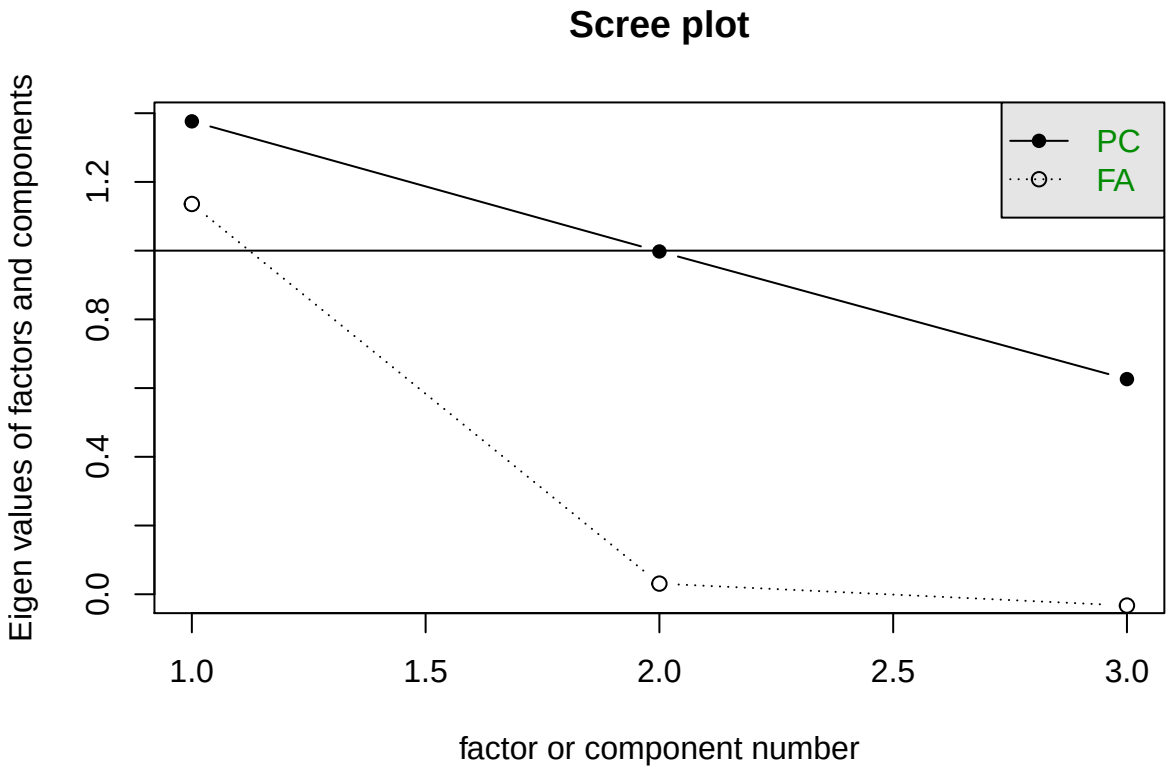
I now run a panel-data regression with my executive-parties dimension (disproportionality, dpi_gf and dpi_enep) as the independent variable and Gini coefficient (wdi_gini) as the dependent variable, controlling for HDI and population size (undp_hdi, unna_pop).

```
# 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")
#   )
```