Third Pair Assignment

19 Apr 2016

Research Question

Our Collaborative Research Paper aims at shedding some light on what the effects of ethnic fractionalization on Wolgang Merkel's conceptualization of political rights are. In this sense, we have two main hypotheses:

H1: The more ethnically fractionalized a country is, the higher its negative impact on political rights if income inequality is high.

H2: When income inequality is low, the impact of high ethnic fractionalization on political rights is expected to be positive.

For more information regarding the theoretical underpinnings of our research proposal, you can have a look at our second pair assignment, which can be found here: Assignment 2

Model

Our aim is to collect data on level of political rights, ethnic fractionalization, and income inequality for all countries in the world during the period 2002-2014. We make use of Freedom House's freedom of the press indicator, drawn from the Quality of Governance dataset, as a proxy for political rights, our dependent variable (DV). Nevertheless, we also intend to make use of an organizational and associational freedom index as an alternative DV for a robustness check.

Our two main independent variables (IVs) of interest are ethnic fractionalization and income inequality (as an interaction variable). Other control variables that we include in our model are GDP per capita, judiciary independence, governmental accountability and linguistic fractionalization. Again, for more information concerning the theoretical reasons behind this model specification, please sese Assignment 2

Data sources

The variables that we are collecting for this research project come from three different data sources: the Quality of Government Standard Data, the World Development Indicators and the Worldwide Governance Indicators.

- Quality of Government (QoG) Standard Time-Series Data - These are contained in the QoG Standard Time-Series Dataset, which includes data from 1946 to 2015, the unit of analysis being country-year. It compiles approximately 2500 variables, making it the largest dataset developed by the Quality of Government Institute at the University of Gothenburg, Sweden.
- World Development Indicators (WDI) - These are a set of indicators developed by the World Bank, compiling "relevant, high-quality, and internationally comparable statistics about global development and the fight against poverty". They are organized around six main themes, which are world view, people, environment, economy, states and markets, and global links.
- Worldwide Governance Indicators (WGI) - These are also developed by the World Bank. They consist of "six composite indicators of broad dimensions of governance covering over 200 countries since 1996: Voice and Accountability, Political Stability and Absence of Violence/Terrorism, Government Effectiveness, Regulatory Quality, Rule of Law, and Control of Corruption". They are obtained from 31 different perception-based data sources, aiming at enabling meaningful cross-country comparisons over time.

Data gathering

Quality of Government data

To collect the data from the Quality of Government Standard Time-Series Dataset (Quality of Government Institute 2015), we downloaded and imported the .csv file containing the data into R and dropped all variables but those we wish to include in our model for the period 2002-2014. These are freedom of the press and associational and organization freedom, both developed by Freedom House; indicators for ethnic and linguistic fractionalization (Alesina et al. 2003); and the World Economic Forum (WEF)'s judicial independence variable.

World Development Indicators

Using the WDI Package for R, we were able to use the World Bank indicator codes to automatically download the variables that we are interested in from the World Bank's APIs. These are: the Gini coefficient, which meassures income inequality (World Bank code: SI.POV.GINI), and GDP per capita in current US dollars (World Bank code: NY.GDP.PCAP.CD) (The World Bank 2014)

Worldwide Governance Indicators

Unlike in the case of the WDI, there is no API. Therefore, we had to download and import the data into R from a zipped .csv file. Since we were only interested in the Voice and Accountability variable for the years 2002-2014, we dropped the other five variables (Political Stability and Absence of Violence/Terrorism, Government Effectiveness, Regulatory Quality, Rule of Law, and Control of Corruption) as well as the values that the variable Voice and Accountability takes for previous years (1996, 1998, 2000) (Kaufmann, Kraay, and Mastruzzi 2010).

Descriptive and Inferential Statistics

Descriptive Statistics

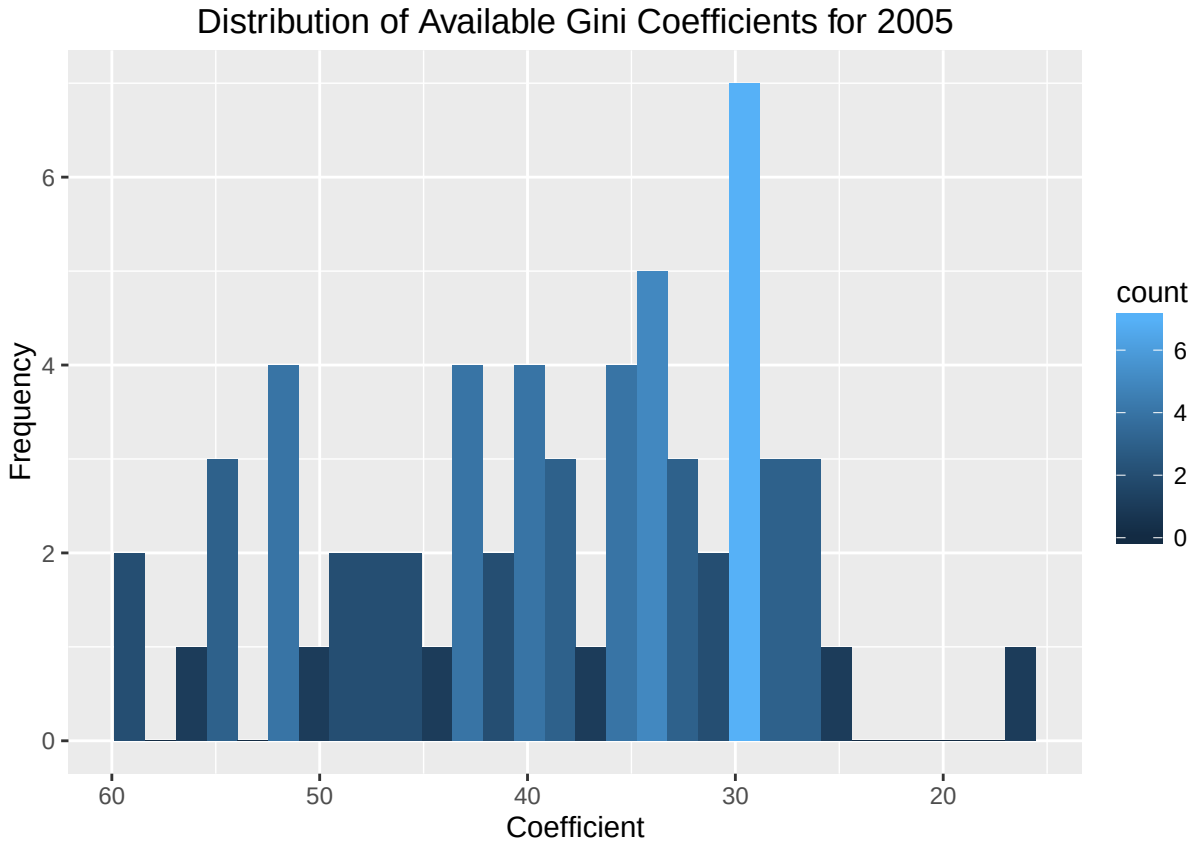
To follow are some basic descriptive statistics examining our main variables of interest: Freedom of the Press (pressfreedom), Ethnic Fractionalization (ethnicfrac), and income inequality as measured by the Gini Coefficient (Ginicoef).

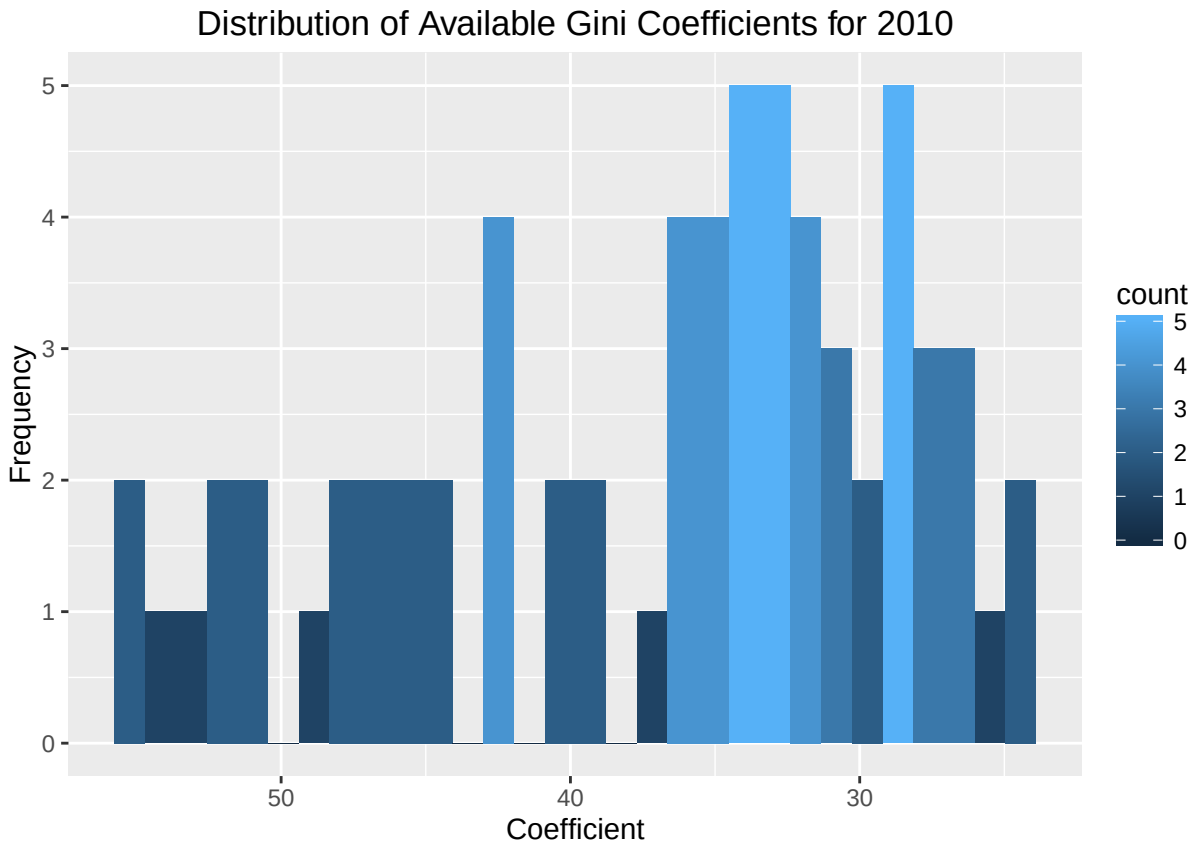
Table 1: Measures of Central Tendency

pressfreedom	ethnicfrac	Ginicoef
Min. : 8.00	Min. :0.0000	Min. :16.64
1st Qu.:23.00	1st Qu.:0.1866	1st Qu.:30.85
Median :44.00	Median :0.4173	Median :35.39
Mean :45.25	Mean :0.4325	Mean :38.10
3rd Qu.:64.00	3rd Qu.:0.6554	3rd Qu.:45.30
Max. :97.00	Max. :0.9302	Max. :64.79
NA	NA	NA's :1641

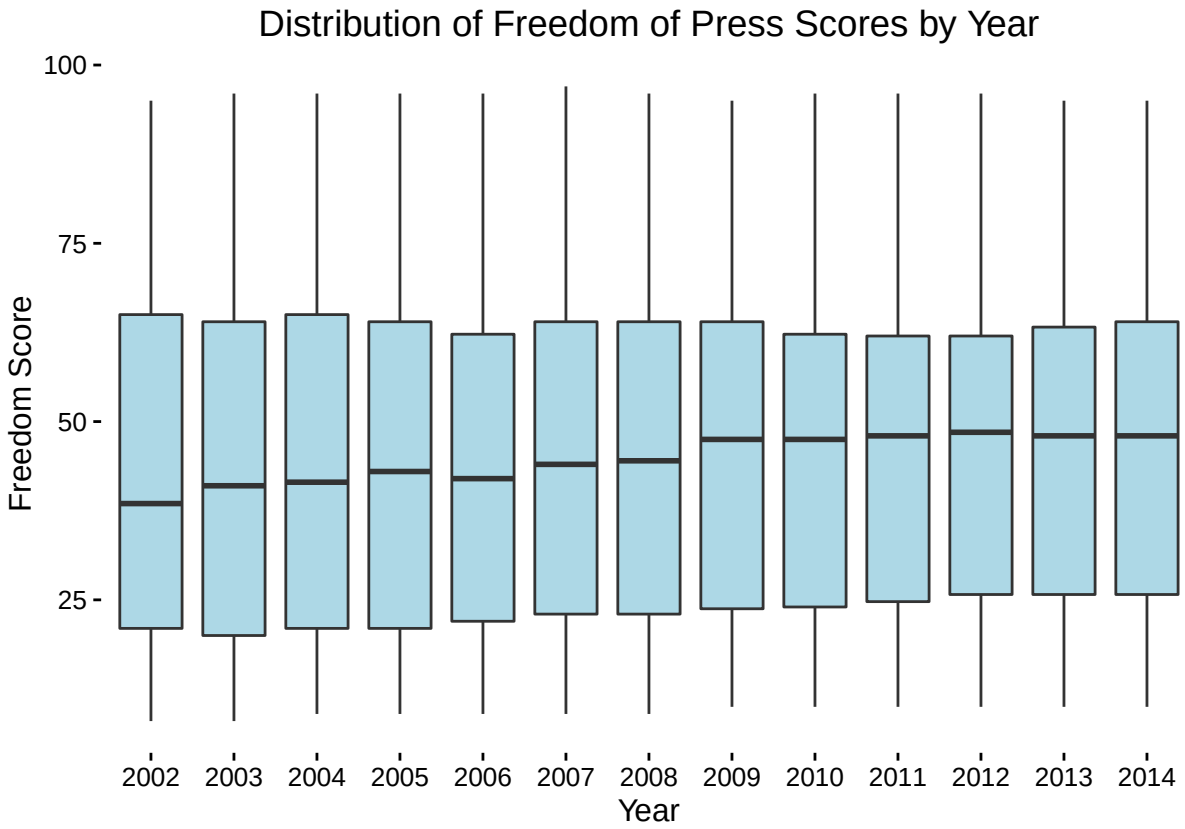
In the following figures, the data has been subset into years to compare change over time. We can see the distribution of countries by relative income inequality for the years 2005 and 2010. These statistics are difficult to interpret, since while we can assume that income inequality does not change so dramatically in most countries over the course of 5 years, the graphs appear drastically different due to

the inconsistency of data available from year to year. In other words, these two charts show different sets of countries; therefore, no conclusion can be drawn as to change in income inequality on a global scale over time.





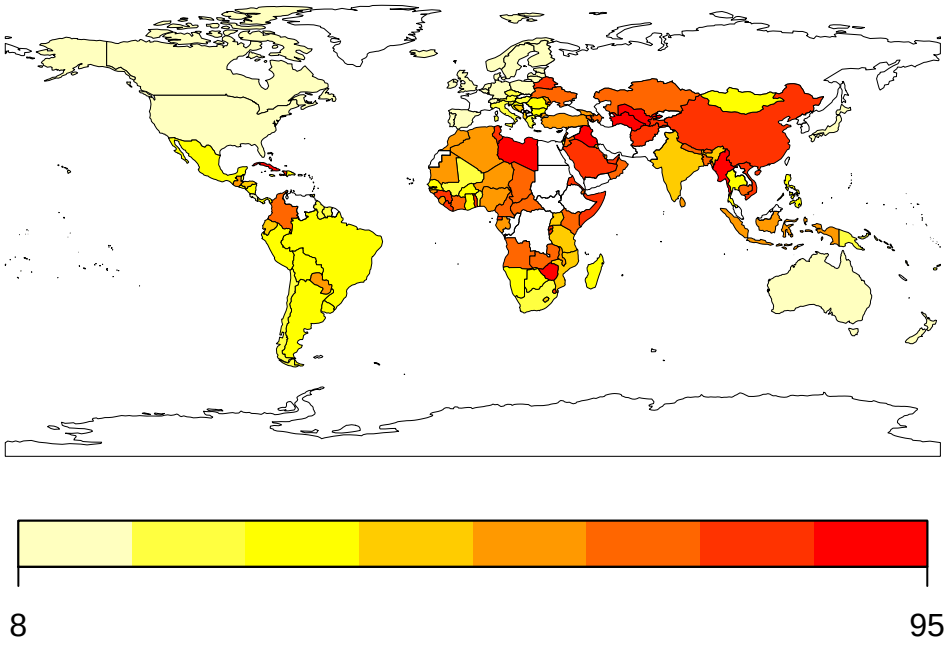
In the figure below, we can see the general distribution of Freedom of Press Scores among the world's countries by year. While the mean score does seem to be slightly increasing over time, there also seems to be a subtle movement from the edges inwards; that is to say, both countries that are on the upper end and the lower end of the spectrum are moving slightly towards the center. It is difficult to detect any significant patterns beyond this, however.



When we inspect press freedom on a map, on the other hand, it is interesting to see how quickly it can fluctuate over time with shifts in political regimes, governments, and societal conditions. The difference here from country to country is much more drastic than one might guess based on the box plot above.

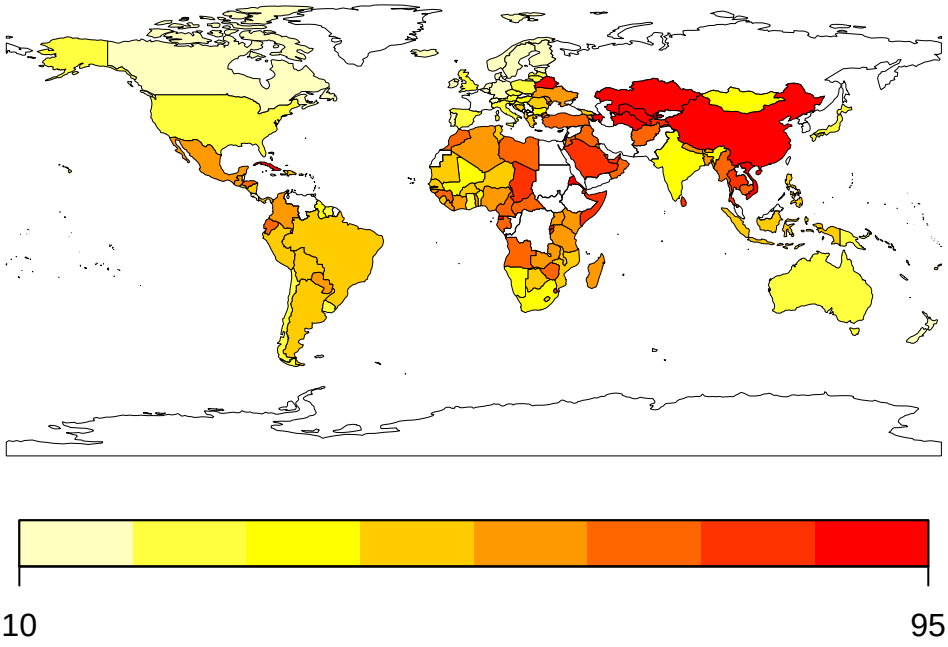
```
## 160 codes from your data successfully matched countries in the map
## 0 codes from your data failed to match with a country code in the map
## 82 codes from the map weren't represented in your data
```

Press Freedom 2002



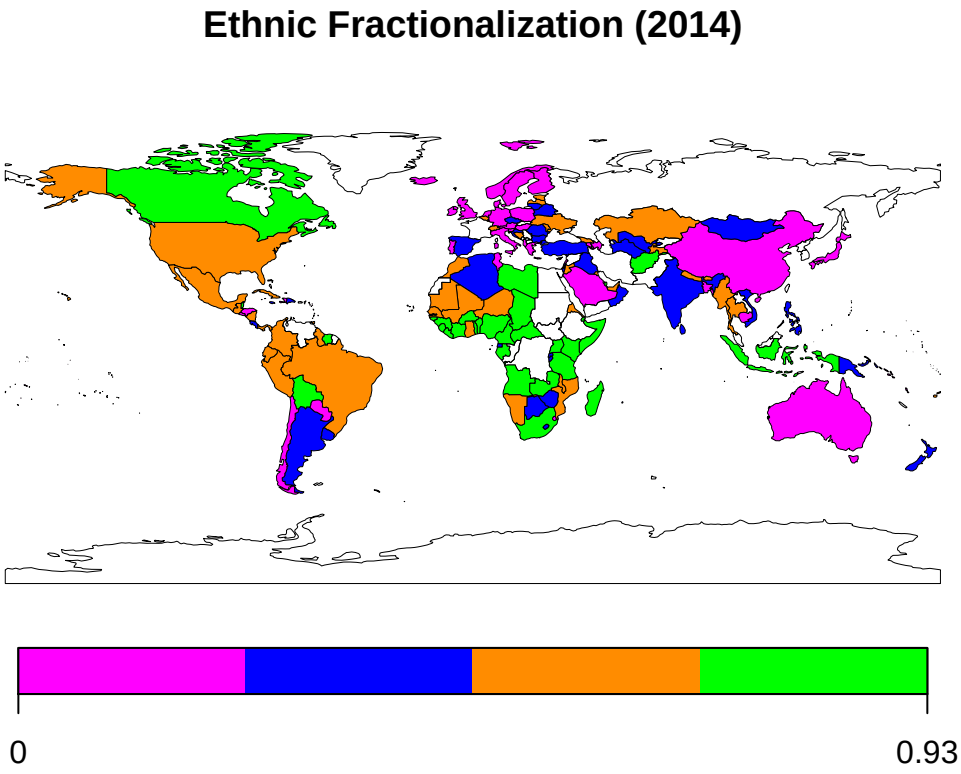

```
## 160 codes from your data successfully matched countries in the map
## 0 codes from your data failed to match with a country code in the map
## 82 codes from the map weren't represented in your data
```

Press Freedom 2014



Since ethnic fractionalization tends to remain relatively steady over time, we have chosen to look at the most recently available statistics. We have categorized countries into bins of relative ethnic

fractionalization, where green is a high probability that any two individuals randomly selected from the population be of different ethnicities, magenta is a low probability (implying high ethnic homogeneity), and blue and orange fall somewhere in the middle. It is interesting to see how the upper half of the scale - those countries in green and orange - and the lower half - blue and magenta - cluster in regions of the world. While we believe this may have implications for similar patterns in strength of political rights, at first glance we see no comparable pattern between this map and the previous ones. The fact that this remains relatively more constant over time than press freedom also suggest that there are likely other factors we need to take into account.



```
##
## Call:
## lm(formula = pressfreedom ~ ethnicfrac + Ginicoef + GDPpercapita +
##     VoiceandAccountability + judindep + langfrac, data = CombinedVars2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -18.9982  -3.3354   0.1988   3.9704  17.7314
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)             4.939e+01  1.971e+00  25.053  < 2e-16 ***
## ethnicfrac             -6.992e+00  1.997e+00  -3.501 0.000526 ***
## Ginicoef                1.168e-01  3.824e-02   3.054 0.002437 **
## GDPpercapita            7.505e-05  2.213e-05   3.392 0.000778 ***
## VoiceandAccountability -2.335e+01  5.979e-01 -39.051  < 2e-16 ***
## judindep               -1.343e+00  3.438e-01  -3.905 0.000114 ***
## langfrac                1.534e+00  1.609e+00   0.953 0.341213
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.413 on 334 degrees of freedom
##   (1739 observations deleted due to missingness)
## Multiple R-squared:  0.9275, Adjusted R-squared:  0.9262
## F-statistic: 712.7 on 6 and 334 DF,  p-value: < 2.2e-16

##
## Call:
## lm(formula = orgassfreedom ~ ethnicfrac + Ginicoef + GDPpercapita +
##     VoiceandAccountability + judindep + langfrac, data = CombinedVars2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.14331 -0.48043 -0.05465  0.39335  3.14238
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)             1.000e+01  3.293e-01  30.368  < 2e-16 ***
## ethnicfrac              6.049e-01  3.336e-01   1.813   0.0707 .
## Ginicoef                1.382e-03  6.388e-03   0.216   0.8289
## GDPpercapita           -2.291e-05  3.696e-06  -6.199 1.68e-09 ***
## VoiceandAccountability  4.239e+00  9.987e-02  42.445  < 2e-16 ***
## judindep               -4.988e-01  5.742e-02  -8.687  < 2e-16 ***
## langfrac                4.511e-02  2.688e-01   0.168   0.8668
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9041 on 334 degrees of freedom
##   (1739 observations deleted due to missingness)
## Multiple R-squared:  0.8943, Adjusted R-squared:  0.8924
## F-statistic: 470.9 on 6 and 334 DF,  p-value: < 2.2e-16

##
## \begin{table}[!htbp] \centering
```

```
##   \caption{OLS regressions}
##   \label{}
## \begin{tabular}{@{\extracolsep{5pt}}lcc}
## \\[-1.8ex]\hline
## \hline \\[-1.8ex]
##  & \multicolumn{2}{c}{\textit{Dependent variable:}} \\
## \cline{2-3}
## \\[-1.8ex] & pressfreedom & orgassfreedom \\
## \\[-1.8ex] & (1) & (2)\\
## \hline \\[-1.8ex]
##  ethnicfrac & $-$6.99$^{***}$ & 0.60$^{*}$ \\
##   & (2.00) & (0.33) \\
##   & & \\
##  Ginicoef & 0.12$^{***}$ & 0.001 \\
##   & (0.04) & (0.01) \\
##   & & \\
##  GDPpercapita & 0.0001$^{***}$ & $-$0.0000$^{***}$ \\
##   & (0.0000) & (0.0000) \\
##   & & \\
##  VoiceandAccountability & $-$23.35$^{***}$ & 4.24$^{***}$ \\
##   & (0.60) & (0.10) \\
##   & & \\
##  judindep & $-$1.34$^{***}$ & $-$0.50$^{***}$ \\
##   & (0.34) & (0.06) \\
##   & & \\
##  langfrac & 1.53 & 0.05 \\
##   & (1.61) & (0.27) \\
##   & & \\
##  Constant & 49.39$^{***}$ & 10.00$^{***}$ \\
##   & (1.97) & (0.33) \\
##   & & \\
## \hline \\[-1.8ex]
## Observations & 341 & 341 \\
## R$^{2}$ & 0.93 & 0.89 \\
## Adjusted R$^{2}$ & 0.93 & 0.89 \\
## Residual Std. Error (df = 334) & 5.41 & 0.90 \\
## F Statistic (df = 6; 334) & 712.65$^{***}$ & 470.85$^{***}$ \\
## \hline
## \hline \\[-1.8ex]
## \textit{Note:}  & \multicolumn{2}{r}{$^{*}$p$<$0.1; $^{**}$p$<$0.05; $^{***}$p$<$0.01} \\
## \end{tabular}
## \end{table}
```

Bibliography

Alesina, Alberto, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat, and Romain Wacziarg. 2003. "Fractionalization." http://www.nber.org/papers/w9411.

Kaufmann, Daniel, Aart Kraay, and Massimo Mastruzzi. 2010. "The Worldwide Governance Indicators: Methodological and Analytical Issues." http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1682130.

Quality of Government Institute. 2015. "QoG Standard Data." http://qog.pol.gu.se/data/datadownloads/qogstandarddata.

The World Bank. 2014. "World Development Indicators." http://data.worldbank.org/sites/default/files/wdi-2014-book.pdf.