

Tutorial 7: Dummy Variables and Interactions (II)

Jan Vogler (jan.vogler@duke.edu)

October 14, 2016

Today's Agenda

1. Fixed effects
2. Interaction terms with interval-level variables
3. Graphical representation of interactions
4. Analysis of Variance (ANOVA)
5. Developing your research proposal
6. Finding variables that represent theoretical concepts

1. Fixed effects

Last time we ran a (pooled) regression and did not include unit-fixed effects. Fixed effects are dummy variables for units or time periods in our data. In the case of the LDC dataset, we can include country fixed effects (and potentially also year or decade fixed effects).

How would we include fixed effects when we run a regression in R?

```
setwd("C:/Users/Jan/OneDrive/Documents/GitHub/ps630_lab/ps630_f16")
library(foreign)
LDC = read.dta("LDC_IO_replication.dta")
main_fe = lm(newtar ~ l1polity + l1signed + l1office + l1gdp_pc + l1lnpop +
             l1ecris2 + l1bpc1 + l1avnewtar + factor(ctylabel) - 1, data = LDC)
summary(main_fe)
```

```
##
## Call:
## lm(formula = newtar ~ l1polity + l1signed + l1office + l1gdp_pc +
##      l1lnpop + l1ecris2 + l1bpc1 + l1avnewtar + factor(ctylabel) -
##      1, data = LDC)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -33.852  -3.563  -0.226   3.311  34.980
##
## Coefficients:
##                                Estimate Std. Error t value
## l1polity                      -5.715e-01  9.872e-02  -5.789
## l1signed                       6.267e-01  7.475e-01   0.838
## l1office                      -2.010e-01  6.881e-02  -2.921
## l1gdp_pc                      -1.476e-04  3.976e-04  -0.371
## l1lnpop                       -1.741e+01  3.414e+00  -5.099
## l1ecris2                      -1.391e+00  1.127e+00  -1.234
## l1bpc1                        2.283e+00  8.801e-01   2.594
## l1avnewtar                     3.739e-01  6.787e-02   5.509
## factor(ctylabel)Albania        2.723e+02  5.217e+01   5.220
```

## factor(ctylabel)Algeria	3.100e+02	5.877e+01	5.276
## factor(ctylabel)Argentina	3.168e+02	5.956e+01	5.318
## factor(ctylabel)Bahrain	2.207e+02	4.575e+01	4.824
## factor(ctylabel)Bangladesh	3.782e+02	6.381e+01	5.927
## factor(ctylabel)Belarus	2.840e+02	5.589e+01	5.081
## factor(ctylabel)Benin	2.926e+02	5.307e+01	5.513
## factor(ctylabel)Bolivia	2.838e+02	5.401e+01	5.254
## factor(ctylabel)Botswana	2.695e+02	4.882e+01	5.520
## factor(ctylabel)Brazil	3.543e+02	6.457e+01	5.486
## factor(ctylabel)Bulgaria	2.925e+02	5.476e+01	5.341
## factor(ctylabel)Burundi	2.910e+02	5.370e+01	5.418
## factor(ctylabel)Cambodia	3.095e+02	5.613e+01	5.514
## factor(ctylabel)Cameroon	2.962e+02	5.602e+01	5.288
## factor(ctylabel)CentralAfricanRepublic	2.734e+02	5.157e+01	5.302
## factor(ctylabel)Chad	2.830e+02	5.449e+01	5.193
## factor(ctylabel)Chile	2.954e+02	5.652e+01	5.227
## factor(ctylabel)China	3.853e+02	7.209e+01	5.345
## factor(ctylabel)Colombia	3.220e+02	5.971e+01	5.393
## factor(ctylabel)Congo	2.695e+02	5.070e+01	5.315
## factor(ctylabel)CostaRica	2.715e+02	5.129e+01	5.293
## factor(ctylabel)Coted'Ivoire	2.958e+02	5.594e+01	5.288
## factor(ctylabel)Cyprus	2.459e+02	4.682e+01	5.253
## factor(ctylabel)DominicanRepublic	2.892e+02	5.449e+01	5.308
## factor(ctylabel)Ecuador	2.985e+02	5.553e+01	5.376
## factor(ctylabel)Egypt	3.362e+02	6.136e+01	5.479
## factor(ctylabel)ElSalvador	2.791e+02	5.329e+01	5.237
## factor(ctylabel)Estonia	2.470e+02	4.900e+01	5.040
## factor(ctylabel)Ethiopia	3.226e+02	6.157e+01	5.241
## factor(ctylabel)Fiji	2.462e+02	4.719e+01	5.217
## factor(ctylabel)Gabon	2.579e+02	4.807e+01	5.366
## factor(ctylabel)Gambia	2.470e+02	4.855e+01	5.087
## factor(ctylabel)Ghana	2.954e+02	5.701e+01	5.181
## factor(ctylabel)Guatemala	2.877e+02	5.526e+01	5.206
## factor(ctylabel)Guinea	2.790e+02	5.417e+01	5.150
## factor(ctylabel)Guyana	2.464e+02	4.712e+01	5.230
## factor(ctylabel)Haiti	2.825e+02	5.413e+01	5.219
## factor(ctylabel)Honduras	2.743e+02	5.379e+01	5.100
## factor(ctylabel)Hungary	2.923e+02	5.567e+01	5.250
## factor(ctylabel)India	4.156e+02	7.056e+01	5.890
## factor(ctylabel)Indonesia	3.435e+02	6.556e+01	5.241
## factor(ctylabel)Israel	2.720e+02	5.307e+01	5.126
## factor(ctylabel)Jamaica	2.681e+02	5.042e+01	5.317
## factor(ctylabel)Jordan	2.659e+02	5.131e+01	5.183
## factor(ctylabel)Kenya	3.163e+02	5.851e+01	5.406
## factor(ctylabel)Korea	3.137e+02	6.062e+01	5.175
## factor(ctylabel)Kuwait	2.404e+02	5.046e+01	4.764
## factor(ctylabel)Latvia	2.599e+02	5.061e+01	5.136
## factor(ctylabel)Lesotho	2.662e+02	5.019e+01	5.303
## factor(ctylabel)Lithuania	2.664e+02	5.184e+01	5.140
## factor(ctylabel)Madagascar	2.844e+02	5.619e+01	5.062
## factor(ctylabel)Malawi	2.878e+02	5.486e+01	5.246
## factor(ctylabel)Malaysia	2.996e+02	5.754e+01	5.207
## factor(ctylabel)Mali	2.954e+02	5.526e+01	5.345
## factor(ctylabel)Mauritania	2.604e+02	5.030e+01	5.178

## factor(ctylabel)Mauritius	2.712e+02	4.769e+01	5.687
## factor(ctylabel)Mexico	3.238e+02	6.281e+01	5.155
## factor(ctylabel)Mongolia	2.609e+02	5.055e+01	5.161
## factor(ctylabel)Morocco	3.136e+02	5.859e+01	5.353
## factor(ctylabel)Mozambique	2.989e+02	5.694e+01	5.249
## factor(ctylabel)Namibia	2.701e+02	4.934e+01	5.474
## factor(ctylabel)Nepal	3.022e+02	5.782e+01	5.226
## factor(ctylabel)Nicaragua	2.718e+02	5.242e+01	5.185
## factor(ctylabel)Niger	2.941e+02	5.535e+01	5.312
## factor(ctylabel)Nigeria	3.406e+02	6.343e+01	5.369
## factor(ctylabel)Oman	2.494e+02	5.040e+01	4.947
## factor(ctylabel)Pakistan	3.761e+02	6.357e+01	5.916
## factor(ctylabel)Panama	2.662e+02	5.079e+01	5.241
## factor(ctylabel)PapuaNewGuinea	2.805e+02	5.230e+01	5.364
## factor(ctylabel)Paraguay	2.721e+02	5.252e+01	5.180
## factor(ctylabel)Peru	3.166e+02	5.799e+01	5.459
## factor(ctylabel)Philippines	3.300e+02	6.153e+01	5.363
## factor(ctylabel)Poland	3.130e+02	5.990e+01	5.226
## factor(ctylabel)Romania	3.029e+02	5.837e+01	5.190
## factor(ctylabel)Russia	3.353e+02	6.454e+01	5.195
## factor(ctylabel)Rwanda	2.995e+02	5.465e+01	5.479
## factor(ctylabel)SaudiArabia	2.835e+02	5.720e+01	4.956
## factor(ctylabel)Senegal	2.781e+02	5.453e+01	5.101
## factor(ctylabel)SierraLeone	2.797e+02	5.267e+01	5.311
## factor(ctylabel)Singapore	2.564e+02	5.233e+01	4.899
## factor(ctylabel)Slovenia	2.625e+02	5.033e+01	5.216
## factor(ctylabel)SouthAfrica	3.096e+02	5.979e+01	5.178
## factor(ctylabel)SriLanka	3.101e+02	5.727e+01	5.414
## factor(ctylabel)Swaziland	2.435e+02	4.813e+01	5.059
## factor(ctylabel)Syria	2.810e+02	5.608e+01	5.011
## factor(ctylabel)Tanzania	3.144e+02	5.915e+01	5.315
## factor(ctylabel)Thailand	3.391e+02	6.133e+01	5.530
## factor(ctylabel)Togo	2.718e+02	5.261e+01	5.166
## factor(ctylabel)Trinidad&Tobago	2.563e+02	4.824e+01	5.312
## factor(ctylabel)Tunisia	2.923e+02	5.483e+01	5.331
## factor(ctylabel)Turkey	3.272e+02	6.121e+01	5.345
## factor(ctylabel)Uganda	2.971e+02	5.771e+01	5.147
## factor(ctylabel)Ukraine	3.154e+02	6.088e+01	5.180
## factor(ctylabel)Uruguay	2.802e+02	5.151e+01	5.440
## factor(ctylabel)Venezuela	3.106e+02	5.768e+01	5.384
## factor(ctylabel)Zambia	2.912e+02	5.488e+01	5.305
## factor(ctylabel)Zimbabwe	2.859e+02	5.563e+01	5.139
##	Pr(> t)		
## l1polity	1.15e-08	***	
## l1signed	0.40213		
## l1office	0.00362	**	
## l1gdp_pc	0.71061		
## l1lnpop	4.61e-07	***	
## l1ecris2	0.21754		
## l1bpc1	0.00971	**	
## l1avnewtar	5.38e-08	***	
## factor(ctylabel)Albania	2.48e-07	***	
## factor(ctylabel)Algeria	1.86e-07	***	
## factor(ctylabel)Argentina	1.49e-07	***	

## factor(ctylabel)Bahrain	1.80e-06	***
## factor(ctylabel)Bangladesh	5.25e-09	***
## factor(ctylabel)Belarus	5.05e-07	***
## factor(ctylabel)Benin	5.27e-08	***
## factor(ctylabel)Bolivia	2.08e-07	***
## factor(ctylabel)Botswana	5.09e-08	***
## factor(ctylabel)Brazil	6.10e-08	***
## factor(ctylabel)Bulgaria	1.32e-07	***
## factor(ctylabel)Burundi	8.79e-08	***
## factor(ctylabel)Cambodia	5.25e-08	***
## factor(ctylabel)Cameroon	1.74e-07	***
## factor(ctylabel)CentralAfricanRepublic	1.62e-07	***
## factor(ctylabel)Chad	2.85e-07	***
## factor(ctylabel)Chile	2.39e-07	***
## factor(ctylabel)China	1.30e-07	***
## factor(ctylabel)Colombia	1.01e-07	***
## factor(ctylabel)Congo	1.51e-07	***
## factor(ctylabel)CostaRica	1.70e-07	***
## factor(ctylabel)Coted'Ivoire	1.74e-07	***
## factor(ctylabel)Cyprus	2.10e-07	***
## factor(ctylabel)DominicanRepublic	1.57e-07	***
## factor(ctylabel)Ecuador	1.10e-07	***
## factor(ctylabel)Egypt	6.34e-08	***
## factor(ctylabel)ElSalvador	2.27e-07	***
## factor(ctylabel)Estonia	6.21e-07	***
## factor(ctylabel)Ethiopia	2.23e-07	***
## factor(ctylabel)Fiji	2.52e-07	***
## factor(ctylabel)Gabon	1.16e-07	***
## factor(ctylabel)Gambia	4.90e-07	***
## factor(ctylabel)Ghana	3.03e-07	***
## factor(ctylabel)Guatemala	2.67e-07	***
## factor(ctylabel)Guinea	3.55e-07	***
## factor(ctylabel)Guyana	2.36e-07	***
## factor(ctylabel)Haiti	2.50e-07	***
## factor(ctylabel)Honduras	4.58e-07	***
## factor(ctylabel)Hungary	2.13e-07	***
## factor(ctylabel)India	6.50e-09	***
## factor(ctylabel)Indonesia	2.23e-07	***
## factor(ctylabel)Israel	4.02e-07	***
## factor(ctylabel)Jamaica	1.50e-07	***
## factor(ctylabel)Jordan	3.01e-07	***
## factor(ctylabel)Kenya	9.36e-08	***
## factor(ctylabel)Korea	3.13e-07	***
## factor(ctylabel)Kuwait	2.40e-06	***
## factor(ctylabel)Latvia	3.83e-07	***
## factor(ctylabel)Lesotho	1.61e-07	***
## factor(ctylabel)Lithuania	3.74e-07	***
## factor(ctylabel)Madagascar	5.56e-07	***
## factor(ctylabel)Malawi	2.17e-07	***
## factor(ctylabel)Malaysia	2.65e-07	***
## factor(ctylabel)Mali	1.29e-07	***
## factor(ctylabel)Mauritania	3.09e-07	***
## factor(ctylabel)Mauritius	2.04e-08	***
## factor(ctylabel)Mexico	3.47e-07	***

```

## factor(ctylabel)Mongolia          3.36e-07 ***
## factor(ctylabel)Morocco           1.24e-07 ***
## factor(ctylabel)Mozambique        2.13e-07 ***
## factor(ctylabel)Namibia           6.53e-08 ***
## factor(ctylabel)Nepal             2.40e-07 ***
## factor(ctylabel)Nicaragua         2.98e-07 ***
## factor(ctylabel)Niger             1.54e-07 ***
## factor(ctylabel)Nigeria          1.14e-07 ***
## factor(ctylabel)Oman              9.84e-07 ***
## factor(ctylabel)Pakistan          5.61e-09 ***
## factor(ctylabel)Panama            2.23e-07 ***
## factor(ctylabel)PapuaNewGuinea    1.17e-07 ***
## factor(ctylabel)Paraguay          3.05e-07 ***
## factor(ctylabel)Peru              7.08e-08 ***
## factor(ctylabel)Philippines       1.17e-07 ***
## factor(ctylabel)Poland            2.40e-07 ***
## factor(ctylabel)Romania           2.90e-07 ***
## factor(ctylabel)Russia            2.82e-07 ***
## factor(ctylabel)Rwanda            6.33e-08 ***
## factor(ctylabel)SaudiArabia       9.44e-07 ***
## factor(ctylabel)Senegal           4.57e-07 ***
## factor(ctylabel)SierraLeone       1.55e-07 ***
## factor(ctylabel)Singapore         1.25e-06 ***
## factor(ctylabel)Slovenia          2.54e-07 ***
## factor(ctylabel)SouthAfrica       3.08e-07 ***
## factor(ctylabel)SriLanka          9.00e-08 ***
## factor(ctylabel)Swaziland         5.64e-07 ***
## factor(ctylabel)Syria             7.18e-07 ***
## factor(ctylabel)Tanzania          1.51e-07 ***
## factor(ctylabel)Thailand          4.83e-08 ***
## factor(ctylabel)Togo              3.27e-07 ***
## factor(ctylabel)Trinidad&Tobago   1.54e-07 ***
## factor(ctylabel)Tunisia           1.40e-07 ***
## factor(ctylabel)Turkey            1.30e-07 ***
## factor(ctylabel)Uganda            3.61e-07 ***
## factor(ctylabel)Ukraine           3.05e-07 ***
## factor(ctylabel)Uruguay           7.82e-08 ***
## factor(ctylabel)Venezuela         1.05e-07 ***
## factor(ctylabel)Zambia            1.60e-07 ***
## factor(ctylabel)Zimbabwe          3.77e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.626 on 589 degrees of freedom
## (4676 observations deleted due to missingness)
## Multiple R-squared:  0.9331, Adjusted R-squared:  0.9212
## F-statistic: 78.24 on 105 and 589 DF, p-value: < 2.2e-16

```

How is this different from our old method? How does it affect the regression we estimate?

2. Interaction terms with interval-level variables

In today's session we will build upon the last lab and continue to work on interaction effects.

Last time we looked at interactions with binary variables. This time both our variables will be measured at the interval level.

For this purpose we will again use the dataset by Milner and Kubota (2005).

Let us try to test a hypothesis that combines claims from modernization theory and the literature on FDI.

1. Modernization theory claims that as countries develop economically, societies become more complex and people more educated which leads to a process of democratization (Lipset 1959, Boix & Stokes 2003).
2. The literature on FDI claims that FDI is an important driver of economic growth. It can increase economic development by allowing for technology transfers from other countries (Jensen 2003, Damijan et al. 2003).

Following these two claims of the political-economic literature, one might argue that there is a positive interaction of GDP and FDI inflows: because FDI means technology transfers—and potentially educational effects—meaning that societies acquire production processes of higher complexity, the effect of economic development on democratization may be magnified.

Let us test this hypothesis through a simple regression model with an interaction term. We use our data on developing countries as there is diversity in terms of all three variables we are interested in: democracy scores, FDI inflows, and GDP per capita.

Let us first specifically look at the values of the variables that we are interested in.

```
summary(LDC$polityiv_update2)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's  
## -10.000  -7.000  -6.000  -2.074   6.000  10.000   2003
```

```
summary(LDC$l1gdp_pc)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's  
##         0     442   1266   2888   2999   44160   1823
```

```
summary(LDC$l1fdi)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's  
## -27.2400  0.0269  0.6382  1.7930  1.9900 184.6000   2423
```

It might be interesting to see what the extreme outliers are in those cases.

Let us check this with the following commands:

```
which(LDC$l1gdp_pc > 20000)
```

```
## [1] 1472 1473 1474 1475 1476 1477 1478 1479 1480 1481 1655 1656 1657 1658  
## [15] 1659 1660 1661 1662 1663 1664 1665 1666 1667 1672 1895 1896 1897 1898  
## [29] 1899 1900 1901 1902 1903 1904 1905 1906 1907 1908 2034 2035 2036 2037  
## [43] 2038 2039 2040 2365 2366 2367 2368 2369 2370
```

```
LDC[1472, ]
```

```
##      country ctylabel date gatt_wto_new aclpn bpc1 dopen_wacz2 ecris2
## 1472    443   Kuwait 1971             1    0    NA          NA    NA
##      fdignp gdp_pc_95d l1aclpn l1bpc1 l1ecris2 newtar polityiv_update2
## 1472    NA  32146.39      0    NA      NA      NA      NA      -8
##      signed yrsoffic      usheg  l1usheg l1fiveop l1gdp_pc      avsw
## 1472      0      NA 0.2678981 0.2756129      10.2 31940.31 0.1505376
##      avnewtar  l1avsw l1avnewtar  lnpop l1lnpop l1office
## 1472      0 0.1505376      0 13.58728 13.5198      NA
##      l1partyage2000 l1fdi l1polity l2polity l3polity l1signed milit2 sp2
## 1472      NA      NA      -9      NA      NA      0      0      0
##      pers2 l1milit2 l1sp2 dictator1 l1dictator1 yr70 yr80 l1ssch closedyr
## 1472      0      0      0      2      2      1      0 1.232      NA
##      _spline1 _spline2 _spline3 l1gatt_wto_new
## 1472      NA      NA      NA      1
```

```
# It's Kuwait
```

```
which(LDC$l1fdi > 50)
```

```
## [1] 984 2847 2848
```

```
LDC[2847, ]
```

```
##      country      ctylabel date gatt_wto_new aclpn bpc1 dopen_wacz2
## 2847    642 EquatorialGuinea 1996             0    0    1          NA
##      ecris2  fdignp gdp_pc_95d l1aclpn l1bpc1 l1ecris2 newtar
## 2847      0 184.5647  517.211      0      1      0      NA
##      polityiv_update2 signed yrsoffic  usheg  l1usheg l1fiveop l1gdp_pc
## 2847      -5      0      17 0.26817 0.2627195      13.2 410.9776
##      avsw avnewtar  l1avsw l1avnewtar  lnpop l1lnpop l1office
## 2847 0.6344086 15.04722 0.5913978  16.13231 12.92255 12.89672      16
##      l1partyage2000  l1fdi l1polity l2polity l3polity l1signed milit2
## 2847      9.5 82.97678      -5      -5      -5      0      0
##      sp2 pers2 l1milit2 l1sp2 dictator1 l1dictator1 yr70 yr80 l1ssch
## 2847      0      0      0      0      2      2      0      0      NA
##      closedyr _spline1 _spline2 _spline3 l1gatt_wto_new
## 2847      NA      NA      NA      NA      0
```

```
# It's Equatorial Guinea
```

Now let us estimate a model with an interaction effect.

```
# Second: estimate the model
```

```
intmodel = lm(polityiv_update2 ~ l1gdp_pc + l1fdi + l1gdp_pc * l1fdi, data = LDC)
summary(intmodel)
```

```
##
```

```
## Call:
```

```
## lm(formula = polityiv_update2 ~ l1gdp_pc + l1fdi + l1gdp_pc *
```

```
##      l1fdi, data = LDC)
##
## Residuals:
##      Min        1Q      Median        3Q        Max
## -21.687   -4.959   -2.689    7.206   12.278
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -2.724e+00  1.951e-01 -13.957 < 2e-16 ***
## l1gdp_pc      1.006e-03  1.042e-04   9.653 < 2e-16 ***
## l1fdi        -5.063e-02  3.696e-02  -1.370  0.171
## l1gdp_pc:l1fdi 1.726e-04  3.475e-05   4.967 7.28e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.755 on 2372 degrees of freedom
## (2994 observations deleted due to missingness)
## Multiple R-squared:  0.08849,    Adjusted R-squared:  0.08734
## F-statistic: 76.76 on 3 and 2372 DF,  p-value: < 2.2e-16
```

How would we interpret the interaction of GDP per capita and FDI? Is it problematic that one of the terms is not significant?

Also, what happens if we use the two formulas below?

```
intmodel2 = lm(polityiv_update2 ~ l1gdp_pc * l1fdi, data = LDC)
summary(intmodel2)
```

```
##
## Call:
## lm(formula = polityiv_update2 ~ l1gdp_pc * l1fdi, data = LDC)
##
## Residuals:
##      Min        1Q      Median        3Q        Max
## -21.687   -4.959   -2.689    7.206   12.278
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -2.724e+00  1.951e-01 -13.957 < 2e-16 ***
## l1gdp_pc      1.006e-03  1.042e-04   9.653 < 2e-16 ***
## l1fdi        -5.063e-02  3.696e-02  -1.370  0.171
## l1gdp_pc:l1fdi 1.726e-04  3.475e-05   4.967 7.28e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.755 on 2372 degrees of freedom
## (2994 observations deleted due to missingness)
## Multiple R-squared:  0.08849,    Adjusted R-squared:  0.08734
## F-statistic: 76.76 on 3 and 2372 DF,  p-value: < 2.2e-16
```

```
intmodel3 = lm(polityiv_update2 ~ l1gdp_pc:l1fdi, data = LDC)
summary(intmodel3)
```



```
##
## Call:
## lm(formula = polityiv_update2 ~ l1gdp_pc:l1fdi, data = LDC)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -27.369  -5.539  -2.722   7.443  15.757
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -1.548e+00  1.528e-01  -10.13  <2e-16 ***
## l1gdp_pc:l1fdi  2.505e-04  2.391e-05   10.48  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.915 on 2374 degrees of freedom
## (2994 observations deleted due to missingness)
## Multiple R-squared:  0.0442, Adjusted R-squared:  0.0438
## F-statistic: 109.8 on 1 and 2374 DF,  p-value: < 2.2e-16
```

Which model would be the better choice?

Note that the interpretation of the main coefficient when there is an interaction term is: when there is a 1-unit increase in x_1 , we expect a b_1 -unit increase in y if x_2 is at 0. But sometimes the value 0 is not meaningful for x_2 . How can we solve this issue?

The answer is: through re-centering. Re-centering means that we simply subtract the mean of x_2 from x_2 .

In order to do this, we will first create a new variable:

```
LDC$l1gdp_pc_cent = LDC$l1gdp_pc - mean(LDC$l1gdp_pc, na.rm = T)
LDC$l1fdi_cent = LDC$l1fdi - mean(LDC$l1fdi, na.rm = T)

summary(LDC$l1gdp_pc_cent)
```

```
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.      NA's
## -2888.0 -2446.0 -1622.0      0.0   110.8 41280.0      1823
```

```
summary(LDC$l1fdi_cent)
```

```
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.      NA's
## -29.0300 -1.7660 -1.1550  0.0000  0.1973 182.8000      2423
```

Now we can run a new regression with this variable.

```
intmodel4 = lm(polityiv_update2 ~ l1gdp_pc_cent + l1fdi_cent + l1gdp_pc_cent *
  l1fdi_cent, data = LDC)
summary(intmodel4)
```

```
##
## Call:
## lm(formula = polityiv_update2 ~ l1gdp_pc_cent + l1fdi_cent +
##      l1gdp_pc_cent * l1fdi_cent, data = LDC)
```

```
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -21.687  -4.959  -2.689   7.206  12.278
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    9.855e-01  1.919e-01   5.137 3.03e-07 ***
## l1gdp_pc_cent    1.316e-03  9.095e-05  14.467 < 2e-16 ***
## l1fdi_cent       4.478e-01  8.005e-02   5.594 2.47e-08 ***
## l1gdp_pc_cent:l1fdi_cent 1.726e-04  3.475e-05   4.967 7.28e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.755 on 2372 degrees of freedom
## (2994 observations deleted due to missingness)
## Multiple R-squared:  0.08849,    Adjusted R-squared:  0.08734
## F-statistic: 76.76 on 3 and 2372 DF,  p-value: < 2.2e-16
```

How would we interpret this?

3. Graphical representation of interactions

There are multiple ways how we can plot interaction terms. If we have the following model:

$$Y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_1 x_2 + u$$

We can plot the effects of interactions in the following ways:

1. We can plot variation in marginal effects of x_1 as x_2 changes (and vice versa).
2. We can plot coefficients and confidence intervals around them.
3. We can plot predicted values of Y as x_1 changes at different values of x_2 .

How would we plot the marginal effect of GDP per capita at different values of FDI inflows?

Let us first check out the coefficients of our model.

```
intmodel$coefficients
```

```
##      (Intercept)      l1gdp_pc      l1fdi l1gdp_pc:l1fdi
## -2.7235846254    0.0010063348  -0.0506297183   0.0001726174
```

Now we need to look at possible values of FDI inflows:

```
summary(LDC$fdignp)
```

```
##      Min.  1st Qu.  Median    Mean 3rd Qu.    Max.     NA's
## -27.2400   0.0361   0.6644   1.8960  2.0830 184.6000    2294
```

```
quantile(LDC$fdignp, probs = c(0.1, 0.9), na.rm = T)
```

```
##      10%      90%
## 0.000000 5.202046
```

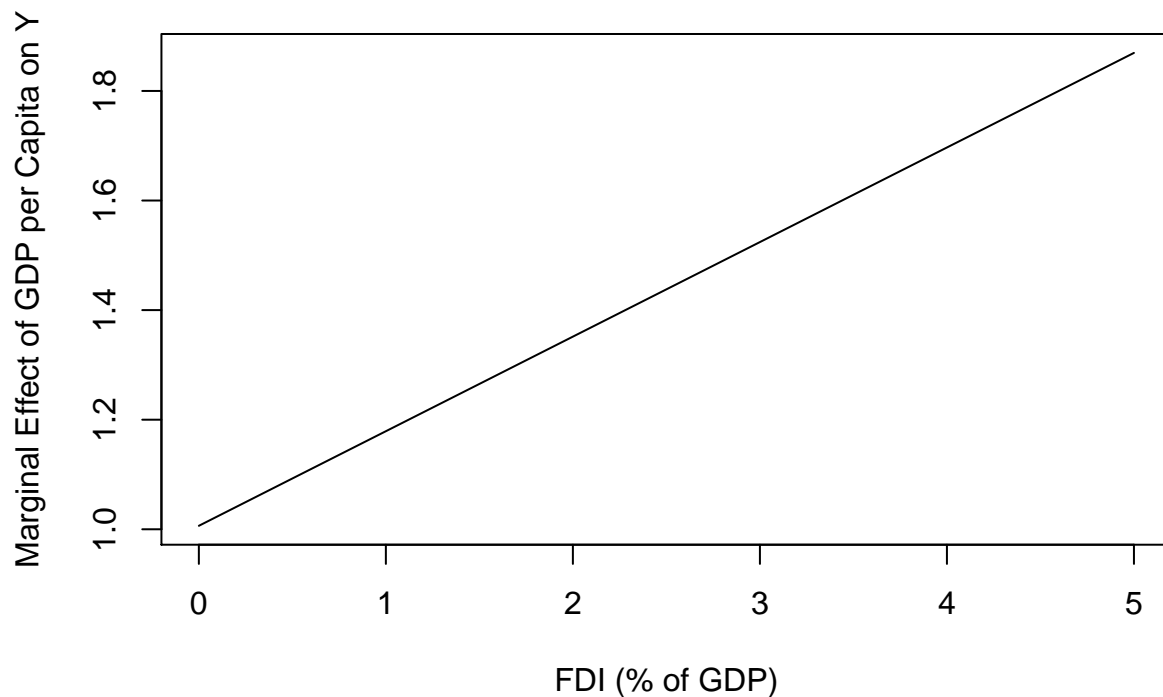
```
fdi_values = seq(0, 5, by = 0.5)
```

```
### Note the marginal effect is for an increase of GDP per Capita by 1 USD In
### order to make interpretation easier, we will look at increases by 1000 USD
```

```
marginal = rep(0.0010063348 * 1000, 11) + 0.0001726174 * fdi_values * 1000
```

```
plot(fdi_values, marginal, type = "l", main = "Marginal Effects of GDP per Capita on Polity IV Score",
     xlab = "FDI (% of GDP)", ylab = "Marginal Effect of GDP per Capita on Y")
```

Marginal Effects of GDP per Capita on Polity IV Score



How would we interpret this plot?

We can easily make a prettier plot that includes confidence intervals by using the package “interplot”:
`install.packages(“interplot”)`

```
library(interplot)
```

```
## Loading required package: ggplot2
```

```
## Loading required package: abind
```

```
## Loading required package: arm
```

```
## Loading required package: MASS
```

```
## Loading required package: Matrix
```

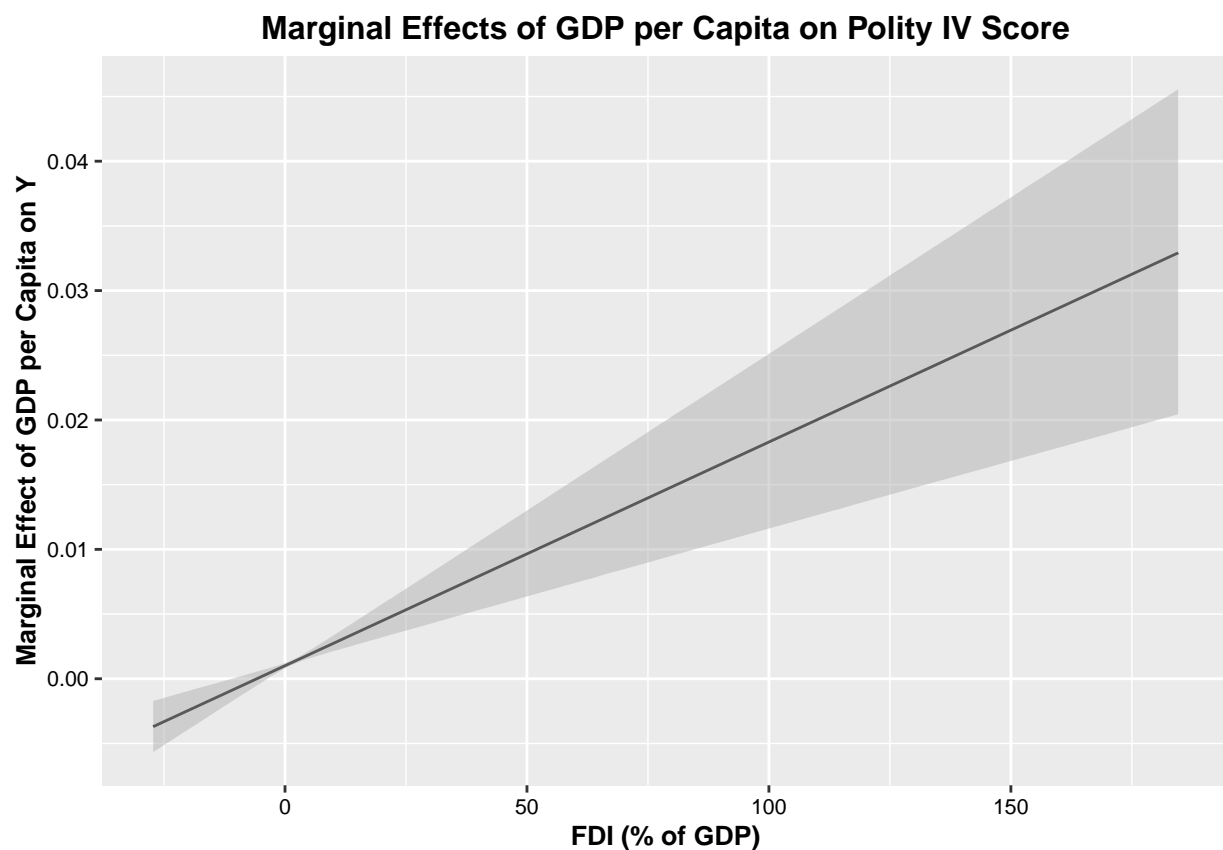
```
## Loading required package: lme4
```

```
##
```

```
## arm (Version 1.9-1, built: 2016-8-21)
```

```
## Working directory is C:/Users/Jan/OneDrive/Documents/GitHub/ps630_lab/ps630_f16/W7
```

```
interplot(m = intmodel, var1 = "l1gdp_pc", var2 = "l1fdi") + xlab("FDI (% of GDP)") +  
  ylab("Marginal Effect of GDP per Capita on Y") + ggtitle("Marginal Effects of GDP per Capita on Pol.  
  theme(plot.title = element_text(face = "bold", size = 12), axis.title = element_text(size = 10,  
    face = "bold"), axis.text = element_text(size = 8, color = "Black"))
```



How would we interpret this plot?

What might we want to do with the data before we run the regression?

Alternatively, we can also plot coefficient plots. These plots show us the coefficients with confidence intervals.

For this purpose, we will use the package “coefplot”: `install.packages(“coefplot”)`

```
library(coefplot)
```

```
##
```

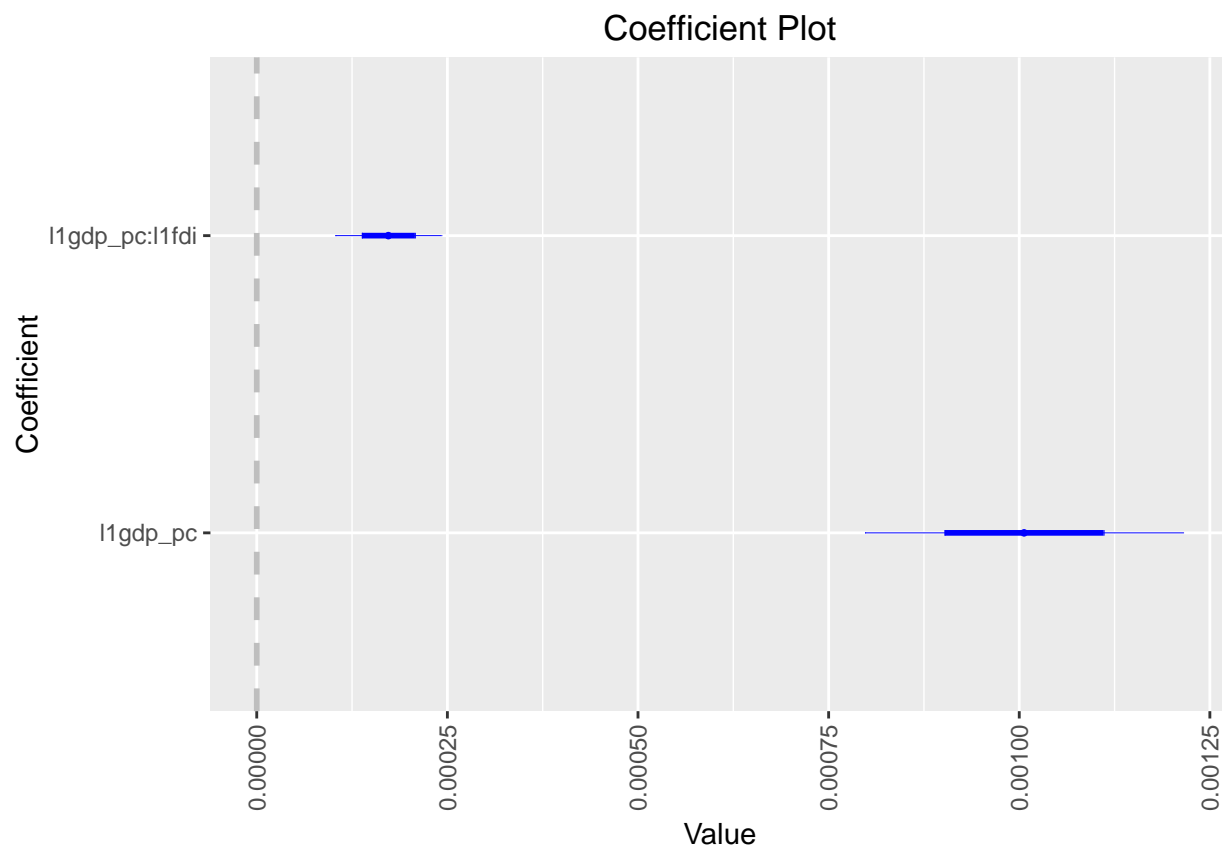
```
## Attaching package: 'coefplot'
```

```
## The following objects are masked from 'package:arm':
##
##   coefplot, coefplot.default
```

```
buildModelCI(intmodel)
```

```
##           Value      Coefficient      HighInner      LowInner
## l1gdp_pc:l1fdi  0.0001726174 l1gdp_pc:l1fdi  0.0002073697  0.0001378651
## l1fdi          -0.0506297183          l1fdi -0.0136661572 -0.0875932793
## l1gdp_pc        0.0010063348          l1gdp_pc  0.0011105892  0.0009020803
## (Intercept)    -2.7235846254    (Intercept) -2.5284411572 -2.9187280936
##           HighOuter      LowOuter      Model
## l1gdp_pc:l1fdi  0.000242122  0.0001031128 intmodel
## l1fdi           0.023297404 -0.1245568404 intmodel
## l1gdp_pc        0.001214844  0.0007978259 intmodel
## (Intercept)    -2.333297689 -3.1138715617 intmodel
```

```
coefplot(intmodel, coefficients = c("l1gdp_pc", "l1gdp_pc:l1fdi"), point = T) +
  theme(axis.text.x = element_text(angle = 90))
```



How would we interpret this plot?

In order to graphically represent the effect that GDP per capita has on our dependent variable (democracy measured by the Polity IV score) at different levels of FDI, we create several new dataframes.

```
quantile(LDC$l1fdi, probs = c(0.25, 0.5, 0.75), na.rm = TRUE)
```

```
##          25%          50%          75%
## 0.0269332 0.6382053 1.9903931
```

```
nd1 = data.frame(l1gdp_pc = seq(1000, 10000, by = 1000), l1fdi = rep(0.0269332,
  10))
nd2 = data.frame(l1gdp_pc = seq(1000, 10000, by = 1000), l1fdi = rep(0.6382053,
  10))
nd3 = data.frame(l1gdp_pc = seq(1000, 10000, by = 1000), l1fdi = rep(1.9903931,
  10))
```

Next we use the model we estimated to predict values based on this new dataframe.

```
pred.p1 = predict(intmodel, type = "response", se.fit = TRUE, newdata = nd1)
pred.p2 = predict(intmodel, type = "response", se.fit = TRUE, newdata = nd2)
pred.p3 = predict(intmodel, type = "response", se.fit = TRUE, newdata = nd3)

pred.table1 = cbind(pred.p1$fit, pred.p1$se.fit)
pred.table2 = cbind(pred.p2$fit, pred.p2$se.fit)
pred.table3 = cbind(pred.p3$fit, pred.p3$se.fit)
```

Finally, we create the plot. Let's start by adding the first values.

```
plot(pred.p1$fit, type = "l", ylim = c(-3, 10), main = "Predicted Values: GDP per capita and Democracy",
  xlab = "GDP per Capita (Thousands of USD)", ylab = "Polity IV Score", axes = FALSE,
  col = "blue", lwd = 2.5)
axis(1, at = seq(1, 10), labels = seq(1, 10))
axis(2, at = seq(-3, 10), labels = seq(-3, 10))

### Next: we add lines

lines(pred.p2$fit, col = "red", lwd = 2.5)
lines(pred.p3$fit, col = "green", lwd = 2.5)

### Let us add a legend to our plot, so it's more obvious what we did here.

legend("bottomright", c("Low FDI", "Median FDI", "High FDI"), lty = 1, lwd = 2,
  col = c("blue", "red", "green"), bty = "n", cex = 1.25)

### We can also add confidence intervals to our plot, though this will make it
### a little less clear.

fit1 = pred.p1$fit
low1 = pred.p1$fit - 2 * pred.p1$se.fit
high1 = pred.p1$fit + 2 * pred.p1$se.fit
cis1 = cbind(fit1, low1, high1)

fit2 = pred.p2$fit
low2 = pred.p2$fit - 2 * pred.p2$se.fit
high2 = pred.p2$fit + 2 * pred.p2$se.fit
cis2 = cbind(fit2, low2, high2)
```

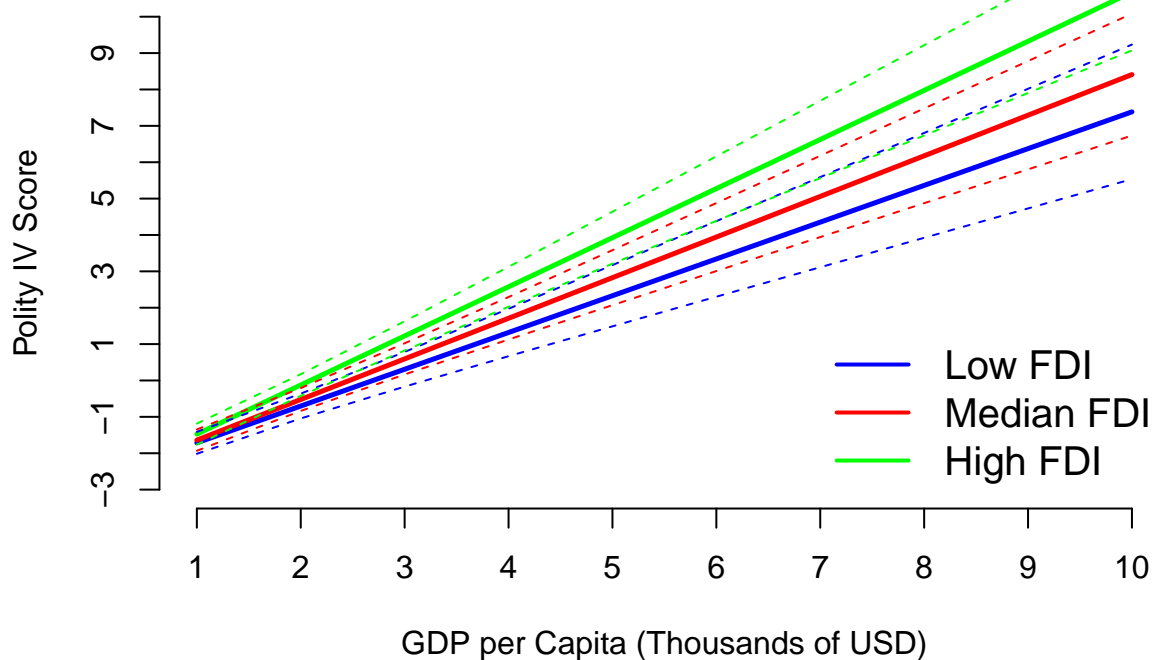
```

fit3 = pred.p3$fit
low3 = pred.p3$fit - 2 * pred.p3$se.fit
high3 = pred.p3$fit + 2 * pred.p3$se.fit
cis3 = cbind(fit3, low3, high3)

matlines(cis1[, c(2, 3)], lty = 2, col = "blue")
matlines(cis2[, c(2, 3)], lty = 2, col = "red")
matlines(cis3[, c(2, 3)], lty = 2, col = "green")

```

Predicted Values: GDP per capita and Democracy Score



How would we interpret this plot?

Be very cautious about any causal interpretation here. We probably have several problems associated with causal inference, including omitted variable bias and endogeneity.

Let's naively assume that we don't deal with those problems here. Would the predicted values support our initial hypothesis?

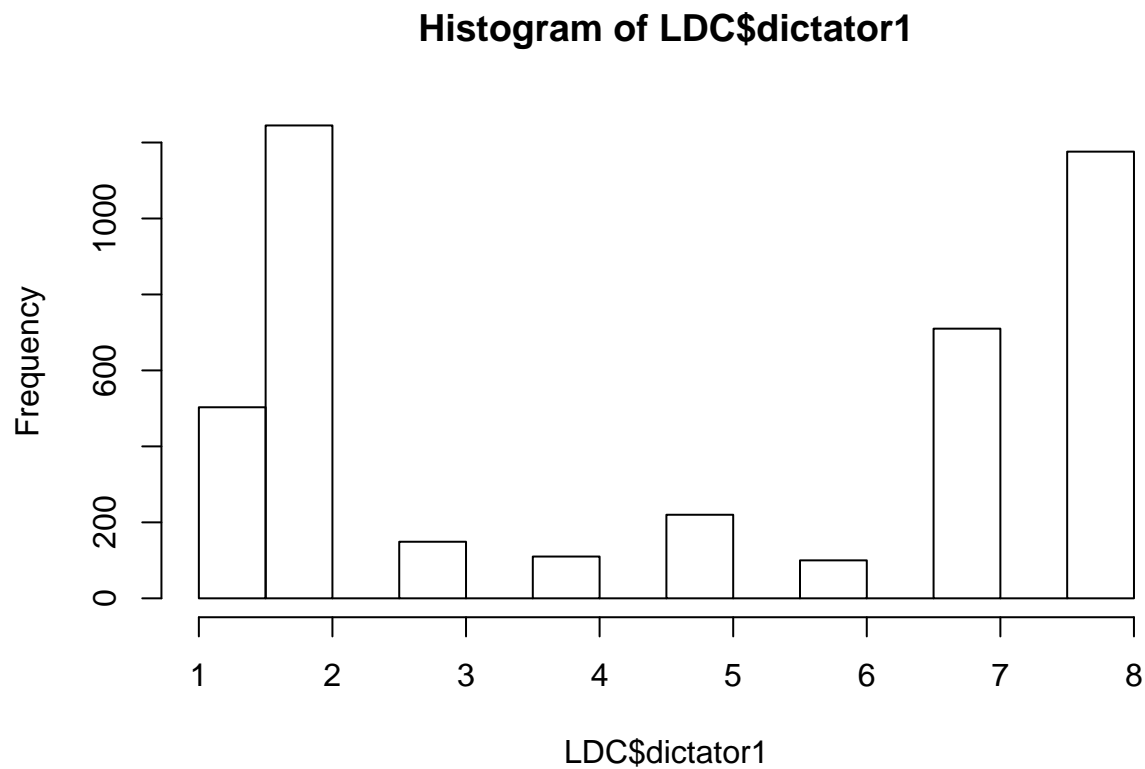
4. Analysis of Variance (ANOVA)

Let us conduct an analysis of variance. We are interested in the level of FDI inflows and we look at one variable:

Type of Dictatorship (dictator1) - coded as values from 1 (personal regime) to 8 (democracy)

Based on Milner & Kubota (2005), how would we expect this variable to influence the level of FDI inflows?

```
hist(LDC$dictator1)
```

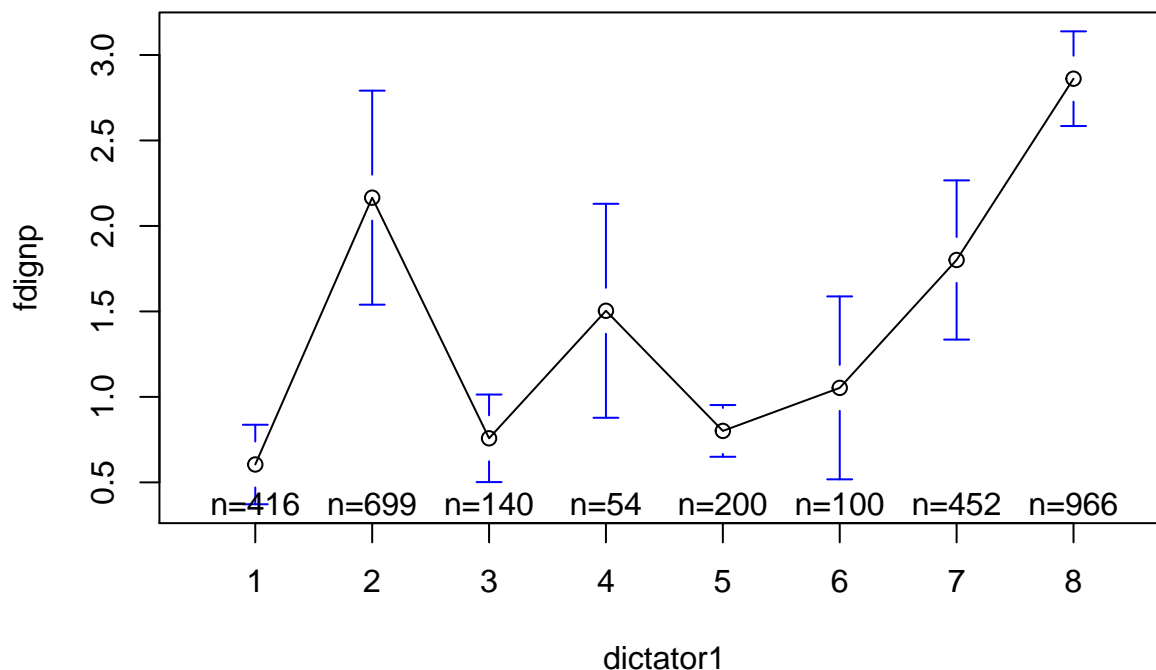


Let us look at a plot of means with the package gplots: `install.packages("gplots")`

```
library(gplots)
```

```
##  
## Attaching package: 'gplots'  
  
## The following object is masked from 'package:stats':  
##  
##     lowess
```

```
plotmeans(fdignp ~ dictator1, data=LDC)
```

summary(LDC)

```
##      country      ctylabel      date      gatt_wto_new
## Min.   :186.0   Length:5370   Min.    :1970   Min.    :0.0000
## 1st Qu.:423.0   Class :character 1st Qu.:1977   1st Qu.:0.0000
## Median :628.0   Mode  :character  Median :1984   Median :0.0000
## Mean   :605.9                      Mean   :1984   Mean   :0.4747
## 3rd Qu.:816.0                      3rd Qu.:1992   3rd Qu.:1.0000
## Max.   :968.0                      Max.    :1999   Max.    :1.0000
##                                     NA's    :698
##      aclpn      bpc1      dopen_wacz2      ecris2
## Min.   :0.0000   Min.    :0.000   Min.    :0.0000   Min.    :0.0000
## 1st Qu.:0.0000   1st Qu.:0.000   1st Qu.:0.0000   1st Qu.:0.0000
## Median :0.0000   Median :1.000   Median :0.0000   Median :0.0000
## Mean   :0.3002   Mean    :0.591   Mean    :0.3097   Mean    :0.0641
## 3rd Qu.:1.0000   3rd Qu.:1.000   3rd Qu.:1.0000   3rd Qu.:0.0000
## Max.   :1.0000   Max.    :1.000   Max.    :1.0000   Max.    :1.0000
## NA's    :1183   NA's    :2734   NA's    :2580   NA's    :1967
##      fdignp      gdp_pc_95d      l1aclpn      l1bpc1
## Min.   :-27.2356   Min.    : 0.0   Min.    :0.0000   Min.    :0.0000
## 1st Qu.: 0.0361   1st Qu.: 442.9   1st Qu.:0.0000   1st Qu.:0.0000
## Median : 0.6644   Median :1266.5   Median :0.0000   Median :1.0000
## Mean    : 1.8962   Mean    :2885.5   Mean    :0.2924   Mean    :0.5909
## 3rd Qu.: 2.0829   3rd Qu.:3002.4   3rd Qu.:1.0000   3rd Qu.:1.0000
## Max.    :184.5647   Max.   :44164.5   Max.    :1.0000   Max.    :1.0000
```

##	NA's :2294	NA's :1679	NA's :1341	NA's :2735
##	llecris2	newtar	polityiv_update2	signed
##	Min. :0.0000	Min. : 0.00	Min. :-10.000	Min. :0.0000
##	1st Qu.:0.0000	1st Qu.: 10.95	1st Qu.: -7.000	1st Qu.:0.0000
##	Median :0.0000	Median : 17.00	Median : -6.000	Median :0.0000
##	Mean :0.0641	Mean : 20.54	Mean : -2.074	Mean :0.1465
##	3rd Qu.:0.0000	3rd Qu.: 27.00	3rd Qu.: 6.000	3rd Qu.:0.0000
##	Max. :1.0000	Max. :102.20	Max. : 10.000	Max. :1.0000
##	NA's :1967	NA's :4463	NA's :2003	NA's :1362
##	yrsoffic	usheg	llusheg	lffiveop
##	Min. : 0.000	Min. :0.2434	Min. :0.2434	Min. :10.20
##	1st Qu.: 2.000	1st Qu.:0.2574	1st Qu.:0.2574	1st Qu.:10.90
##	Median : 5.000	Median :0.2663	Median :0.2655	Median :12.35
##	Mean : 8.431	Mean :0.2696	Mean :0.2683	Mean :12.03
##	3rd Qu.:12.000	3rd Qu.:0.2785	3rd Qu.:0.2784	3rd Qu.:12.72
##	Max. :44.000	Max. :0.3083	Max. :0.2988	Max. :13.20
##	NA's :2361		NA's :179	NA's :358
##	llgdp_pc	avsw	avnewtar	llavsw
##	Min. : 0	Min. :0.1398	Min. : 0.00	Min. :0.1398
##	1st Qu.: 442	1st Qu.:0.1505	1st Qu.: 0.00	1st Qu.:0.1505
##	Median : 1266	Median :0.1720	Median :17.43	Median :0.1613
##	Mean : 2888	Mean :0.3097	Mean :14.91	Mean :0.2974
##	3rd Qu.: 2999	3rd Qu.:0.5269	3rd Qu.:24.37	3rd Qu.:0.5054
##	Max. :44165	Max. :0.6667	Max. :30.52	Max. :0.6559
##	NA's :1823			NA's :179
##	llavnewtar	lnpop	lllnpop	lloffice
##	Min. : 0.00	Min. :10.57	Min. :10.62	Min. : 0.000
##	1st Qu.: 0.00	1st Qu.:13.86	1st Qu.:13.86	1st Qu.: 2.000
##	Median :18.73	Median :15.32	Median :15.31	Median : 5.000
##	Mean :15.01	Mean :15.11	Mean :15.10	Mean : 8.431
##	3rd Qu.:24.37	3rd Qu.:16.40	3rd Qu.:16.39	3rd Qu.:12.000
##	Max. :30.52	Max. :20.95	Max. :20.94	Max. :44.000
##	NA's :179	NA's :490	NA's :661	NA's :2361
##	llpartyage2000	llfdi	llpolity	l2polity
##	Min. : 0.00	Min. : -27.2356	Min. : -10.000	Min. : -10.00
##	1st Qu.: 10.00	1st Qu.: 0.0269	1st Qu.: -7.000	1st Qu.: -7.00
##	Median : 19.50	Median : 0.6382	Median : -6.000	Median : -7.00
##	Mean : 24.18	Mean : 1.7931	Mean : -2.215	Mean : -2.36
##	3rd Qu.: 32.00	3rd Qu.: 1.9904	3rd Qu.: 6.000	3rd Qu.: 5.00
##	Max. :183.00	Max. :184.5647	Max. : 10.000	Max. : 10.00
##	NA's :3284	NA's :2423	NA's :2124	NA's :2246
##	l3polity	llsigned	milit2	sp2
##	Min. : -10.000	Min. :0.0000	Min. :0.0000	Min. :0.0000
##	1st Qu.: -7.000	1st Qu.:0.0000	1st Qu.:0.0000	1st Qu.:0.0000
##	Median : -7.000	Median :0.0000	Median :0.0000	Median :0.0000
##	Mean : -2.512	Mean :0.1511	Mean :0.1119	Mean :0.1959
##	3rd Qu.: 5.000	3rd Qu.:0.0000	3rd Qu.:0.0000	3rd Qu.:0.0000
##	Max. : 10.000	Max. :1.0000	Max. :1.0000	Max. :1.0000
##	NA's :2371	NA's :1517		
##	pers2	llmilit2	llsp2	dictator1
##	Min. :0.0000	Min. :0.0000	Min. :0.0000	Min. :1.000
##	1st Qu.:0.0000	1st Qu.:0.0000	1st Qu.:0.0000	1st Qu.:2.000
##	Median :0.0000	Median :0.0000	Median :0.0000	Median :5.000
##	Mean :0.1665	Mean :0.1135	Mean :0.1986	Mean :4.737

```
## 3rd Qu.:0.0000 3rd Qu.:0.0000 3rd Qu.:0.0000 3rd Qu.:8.000
## Max. :1.0000 Max. :1.0000 Max. :1.0000 Max. :8.000
## NA's :179 NA's :179 NA's :1157
## lldictator1 yr70 yr80 l1ssch
## Min. :1.000 Min. :0.0000 Min. :0.0000 Min. :0.0140
## 1st Qu.:2.000 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.4562
## Median :5.000 Median :0.0000 Median :0.0000 Median :0.8519
## Mean :4.708 Mean :0.3333 Mean :0.3333 Mean :1.0411
## 3rd Qu.:8.000 3rd Qu.:1.0000 3rd Qu.:1.0000 3rd Qu.:1.4652
## Max. :8.000 Max. :1.0000 Max. :1.0000 Max. :4.4422
## NA's :1315 NA's :3140
## closedyr _spline1 _spline2 _spline3
## Min. : 0.000 Min. : -24389 Min. : -7854.0 Min. : -9030.0
## 1st Qu.: 0.000 1st Qu.: -3375 1st Qu.: -2048.3 1st Qu.: -1629.3
## Median : 7.000 Median : -343 Median : -260.2 Median : -165.6
## Mean : 8.691 Mean : -3075 Mean : -1388.8 Mean : -1340.9
## 3rd Qu.:15.000 3rd Qu.: 0 3rd Qu.: 0.0 3rd Qu.: 0.0
## Max. :29.000 Max. : 0 Max. : 0.0 Max. : 0.0
## NA's :2580 NA's :2580 NA's :2580 NA's :2580
## l1gatt_wto_new l1gdp_pc_cent l1fdi_cent
## Min. :0.000 Min. : -2887.8 Min. : -29.0287
## 1st Qu.:0.000 1st Qu.: -2445.8 1st Qu.: -1.7661
## Median :0.000 Median : -1621.5 Median : -1.1549
## Mean :0.468 Mean : 0.0 Mean : 0.0000
## 3rd Qu.:1.000 3rd Qu.: 110.8 3rd Qu.: 0.1973
## Max. :1.000 Max. :41276.7 Max. :182.7717
## NA's :868 NA's :1823 NA's :2423
```

Let us look at the ANOVA results of this variable.

```
anovamodel = lm(fdignp ~ factor(dictator1), data = LDC)
summary(anovamodel)
```

```
##
## Call:
## lm(formula = fdignp ~ factor(dictator1), data = LDC)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -27.840  -1.909   -0.736    0.300  182.399
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.6046     0.2580   2.344 0.019158 *
## factor(dictator1)2  1.5609     0.3258   4.791 1.74e-06 ***
## factor(dictator1)3  0.1532     0.5141    0.298 0.765783
## factor(dictator1)4  0.8991     0.7610    1.181 0.237515
## factor(dictator1)5  0.1964     0.4527    0.434 0.664431
## factor(dictator1)6  0.4482     0.5860    0.765 0.444424
## factor(dictator1)7  1.1963     0.3575    3.347 0.000828 ***
## factor(dictator1)8  2.2568     0.3085    7.314 3.30e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 5.261 on 3019 degrees of freedom
## (2343 observations deleted due to missingness)
## Multiple R-squared: 0.02506, Adjusted R-squared: 0.0228
## F-statistic: 11.09 on 7 and 3019 DF, p-value: 6.499e-14

anova(anovamodel)

## Analysis of Variance Table
##
## Response: fdignp
##              Df Sum Sq Mean Sq F value    Pr(>F)
## factor(dictator1)    7    2148   306.92    11.087 6.499e-14 ***
## Residuals          3019   83575    27.683
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

aov_fdi = aov(fdignp ~ factor(dictator1), data = LDC)
summary(aov_fdi)

##              Df Sum Sq Mean Sq F value    Pr(>F)
## factor(dictator1)    7    2148   306.92    11.09 6.5e-14 ***
## Residuals          3019   83575    27.68
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 2343 observations deleted due to missingness

tukey = TukeyHSD(aov_fdi)
tukey

## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = fdignp ~ factor(dictator1), data = LDC)
##
## $`factor(dictator1)`
##              diff              lwr              upr              p adj
## 2-1  1.56088894  0.57271672  2.54906116  0.0000475
## 3-1  0.15315662 -1.40606258  1.71237582  0.9999899
## 4-1  0.89914163 -1.40912432  3.20740759  0.9373334
## 5-1  0.19641327 -1.17670882  1.56953537  0.9998698
## 6-1  0.44818477 -1.32910667  2.22547620  0.9948035
## 7-1  1.19633535  0.11209780  2.28057289  0.0187714
## 8-1  2.25679265  1.32095772  3.19262759  0.0000000
## 3-2 -1.40773232 -2.88533897  0.06987433  0.0750121
## 4-2 -0.66174731 -2.91568799  1.59219337  0.9869664
## 5-2 -1.36447567 -2.64417140 -0.08477993  0.0270736
## 6-2 -1.11270418 -2.81884653  0.59343818  0.4969781
## 7-2 -0.36455360 -1.32773994  0.59863275  0.9460293
## 8-2  0.69590371 -0.09652541  1.48833284  0.1342884
## 4-3  0.74598501 -1.81036566  3.30233568  0.9874262
## 5-3  0.04325665 -1.71523740  1.80175071  1.0000000
```

```
## 6-3  0.29502814 -1.79437414  2.38443042  0.9998806
## 7-3  1.04317872 -0.50032642  2.58668386  0.4482069
## 8-3  2.10363603  0.66050607  3.54676599  0.0002719
## 5-4 -0.70272836 -3.15001840  1.74456168  0.9885736
## 6-4 -0.45095687 -3.14586486  2.24395112  0.9996285
## 7-4  0.29719371 -2.00048671  2.59487413  0.9999340
## 8-4  1.35765102 -0.87383980  3.58914184  0.5889220
## 6-5  0.25177149 -1.70268538  2.20622836  0.9999358
## 7-5  0.99992207 -0.35532989  2.35517403  0.3292688
## 8-5  2.06037938  0.82065205  3.30010671  0.0000135
## 7-6  0.74815058 -1.01537102  2.51167218  0.9039984
## 8-6  1.80860789  0.13223549  3.48498028  0.0239452
## 8-7  1.06045731  0.15104506  1.96986955  0.0097654
```

How would we interpret these results? Compare to the results of the OLS regression.

5. Developing your research proposal

1. What is your theory?
2. How can you turn this theory into a testable hypothesis?
3. Which data is out there that would allow you to test your hypothesis?

6. Finding variables that represent theoretical concepts

We are interested in finding a measurement for veto players. Veto players can be described as the number of institutions in a political system whose approval is required when changes to the status quo are attempted to be made.

How can we measure this theoretical concept? Which variables might be most useful?

In class-exercise: turning a concept into a measurement