

# Pol Sci 630: Problem Set 9 - Data Management and Omitted Variable Bias - Solutions

Prepared by: Jan Vogler (jan.vogler@duke.edu)

Grading Due Date: Friday, October 30th, 12.00 PM (Beginning of Lab)

**Insert your comments on the assignment that you are grading above the solution in bold and red text. For example write: "GRADER COMMENT: everything is correct! - 4/4 Points" Also briefly point out which, if any, problems were not solved correctly and what the mistake was.**

Use the following scheme to assign points: For problems that were solved correctly in their entirety, assign the full point value of 4. For correctly solved bonus problems, add that value to the total score for a problem but do not go above 4 points per problem. If there are mistakes in any problem, subtract points according to the extent of the mistake. If you subtract points, explain why.

In order to make your text bold and red, you need to insert the following line at the beginning of the document:

```
\usepackage{color}
```

and the following lines above the solution of the specific task:

```
\textbf{\color{red} GRADER COMMENT: everything is correct! - 4/4 Points}
```

# R Programming

## Problem 1

```
#### a
setwd("C:/Users/Jan/OneDrive/Documents/GitHub/ps630_lab/")
library(foreign)
LDC = read.dta("LDC_IO_replication.dta")

LDC$regime_class = NA
LDC$regime_class[LDC$polityiv_update2 > 5] = "Democracy"
unique(LDC$regime_class)

## [1] "Democracy" NA

LDC$regime_class[LDC$polityiv_update2 >= -5 & LDC$polityiv_update2 <= 5] = "Anocracy"
unique(LDC$regime_class)

## [1] "Democracy" "Anocracy" NA

LDC$regime_class[LDC$polityiv_update2 < -5] = "Autocracy"
unique(LDC$regime_class)

## [1] "Democracy" "Anocracy" "Autocracy" NA

#### b

LDC$l1polity_squared = (LDC$l1polity)^2

lm_fdi2 = lm(fdignp ~ l1polity + l1polity_squared + l1signed + l1office + l1gdp_pc +
  l1lnpop + l1ecris2 + l1bpc1 + l1avnewtar + factor(ctylabel) - 1, data = LDC)
summary(lm_fdi2)

##
## Call:
## lm(formula = fdignp ~ l1polity + l1polity_squared + l1signed +
##     l1office + l1gdp_pc + l1lnpop + l1ecris2 + l1bpc1 + l1avnewtar +
```

```
##      factor(ctylabel) - 1, data = LDC)
##
## Residuals:
##      Min        1Q    Median        3Q        Max
## -32.027   -0.908   -0.124    0.604  152.058
##
## Coefficients:
##                                     Estimate Std. Error t value
## l1polity                        4.233e-02  3.407e-02   1.242
## l1polity_squared                8.910e-03  7.833e-03   1.138
## l1signed                       -2.362e-01  3.101e-01  -0.762
## l1office                       -3.138e-02  2.554e-02  -1.229
## l1gdp_pc                       -1.431e-04  3.028e-04  -0.473
## l1lnpop                        3.489e+00  1.049e+00   3.326
## l1ecris2                       8.556e-02  4.716e-01   0.181
## l1bpc1                        -4.308e-01  3.520e-01  -1.224
## l1avnewtar                     -4.951e-02  1.399e-02  -3.540
## factor(ctylabel)Albania        -4.911e+01  1.575e+01  -3.119
## factor(ctylabel)Algeria        -5.773e+01  1.787e+01  -3.231
## factor(ctylabel)Angola         -1.869e+01  1.769e+01  -1.057
## factor(ctylabel)Argentina      -5.749e+01  1.795e+01  -3.203
## factor(ctylabel)Armenia        -4.816e+01  1.602e+01  -3.006
## factor(ctylabel)Azerbaijan     -3.940e+01  1.679e+01  -2.347
## factor(ctylabel)Bangladesh     -6.302e+01  1.939e+01  -3.249
## factor(ctylabel)Belarus        -5.453e+01  1.707e+01  -3.195
## factor(ctylabel)Benin          -5.150e+01  1.609e+01  -3.201
## factor(ctylabel)Bhutan         -4.481e+01  1.412e+01  -3.174
## factor(ctylabel)Bolivia        -5.160e+01  1.645e+01  -3.136
## factor(ctylabel)Botswana       -4.479e+01  1.453e+01  -3.083
## factor(ctylabel)Brazil         -6.306e+01  1.951e+01  -3.233
## factor(ctylabel)Bulgaria       -5.299e+01  1.673e+01  -3.167
## factor(ctylabel)Burundi        -5.264e+01  1.643e+01  -3.203
## factor(ctylabel)Cambodia       -4.879e+01  1.718e+01  -2.841
## factor(ctylabel)Cameroon       -5.387e+01  1.709e+01  -3.152
```

|   |            |           |        |
|---|------------|-----------|--------|
| ## factor(ctylabel)CentralAfricanRepublic | -4.986e+01 | 1.564e+01 | -3.187 |
| ## factor(ctylabel)Chad                   | -5.192e+01 | 1.649e+01 | -3.149 |
| ## factor(ctylabel)Chile                  | -5.327e+01 | 1.707e+01 | -3.121 |
| ## factor(ctylabel)China                  | -6.884e+01 | 2.199e+01 | -3.130 |
| ## factor(ctylabel)Colombia               | -5.843e+01 | 1.804e+01 | -3.240 |
| ## factor(ctylabel)Comoros                | -4.304e+01 | 1.360e+01 | -3.164 |
| ## factor(ctylabel)Congo                  | -4.814e+01 | 1.533e+01 | -3.140 |
| ## factor(ctylabel)CostaRica              | -4.908e+01 | 1.552e+01 | -3.162 |
| ## factor(ctylabel)Coted'Ivoire           | -5.337e+01 | 1.711e+01 | -3.118 |
| ## factor(ctylabel)Croatia                | -4.941e+01 | 1.618e+01 | -3.055 |
| ## factor(ctylabel)DominicanRepublic      | -5.184e+01 | 1.639e+01 | -3.163 |
| ## factor(ctylabel)Ecuador                | -5.403e+01 | 1.680e+01 | -3.216 |
| ## factor(ctylabel)Egypt                  | -5.786e+01 | 1.858e+01 | -3.113 |
| ## factor(ctylabel)ElSalvador             | -5.222e+01 | 1.613e+01 | -3.238 |
| ## factor(ctylabel)EquatorialGuinea       | -1.072e+01 | 1.363e+01 | -0.786 |
| ## factor(ctylabel)Estonia                | -4.228e+01 | 1.488e+01 | -2.842 |
| ## factor(ctylabel)Ethiopia               | -5.969e+01 | 1.868e+01 | -3.195 |
| ## factor(ctylabel)Fiji                   | -4.283e+01 | 1.403e+01 | -3.053 |
| ## factor(ctylabel)Gabon                  | -4.337e+01 | 1.440e+01 | -3.012 |
| ## factor(ctylabel)Gambia                 | -4.447e+01 | 1.428e+01 | -3.114 |
| ## factor(ctylabel)Georgia                | -4.577e+01 | 1.695e+01 | -2.700 |
| ## factor(ctylabel)Ghana                  | -5.517e+01 | 1.726e+01 | -3.196 |
| ## factor(ctylabel)Guatemala              | -5.302e+01 | 1.664e+01 | -3.186 |
| ## factor(ctylabel)Guinea                 | -5.255e+01 | 1.655e+01 | -3.175 |
| ## factor(ctylabel)GuineaBissau           | -4.587e+01 | 1.462e+01 | -3.138 |
| ## factor(ctylabel)Guyana                 | -4.010e+01 | 1.427e+01 | -2.809 |
| ## factor(ctylabel)Haiti                  | -5.275e+01 | 1.657e+01 | -3.184 |
| ## factor(ctylabel)Honduras               | -5.106e+01 | 1.601e+01 | -3.190 |
| ## factor(ctylabel)Hungary                | -5.165e+01 | 1.686e+01 | -3.063 |
| ## factor(ctylabel)India                  | -7.108e+01 | 2.145e+01 | -3.314 |
| ## factor(ctylabel)Indonesia              | -6.326e+01 | 2.000e+01 | -3.163 |
| ## factor(ctylabel)Iran                   | -5.967e+01 | 1.861e+01 | -3.206 |
| ## factor(ctylabel)Jamaica                | -4.903e+01 | 1.539e+01 | -3.185 |
| ## factor(ctylabel)Jordan                 | -4.865e+01 | 1.568e+01 | -3.102 |

|                                   |            |           |        |
|-----------------------------------|------------|-----------|--------|
| ## factor(ctylabel)Kazakhstan     | -5.184e+01 | 1.747e+01 | -2.968 |
| ## factor(ctylabel)Kenya          | -5.656e+01 | 1.778e+01 | -3.181 |
| ## factor(ctylabel)Korea          | -5.861e+01 | 1.827e+01 | -3.208 |
| ## factor(ctylabel)KyrgyzRepublic | -4.754e+01 | 1.623e+01 | -2.930 |
| ## factor(ctylabel)Laos           | -4.850e+01 | 1.626e+01 | -2.982 |
| ## factor(ctylabel)Latvia         | -4.492e+01 | 1.551e+01 | -2.897 |
| ## factor(ctylabel)Lesotho        | -4.394e+01 | 1.513e+01 | -2.904 |
| ## factor(ctylabel)Lithuania      | -4.921e+01 | 1.599e+01 | -3.077 |
| ## factor(ctylabel)Madagascar     | -5.475e+01 | 1.705e+01 | -3.212 |
| ## factor(ctylabel)Malawi         | -5.312e+01 | 1.683e+01 | -3.156 |
| ## factor(ctylabel)Malaysia       | -5.219e+01 | 1.722e+01 | -3.030 |
| ## factor(ctylabel)Mali           | -5.332e+01 | 1.674e+01 | -3.185 |
| ## factor(ctylabel)Mauritania     | -4.752e+01 | 1.529e+01 | -3.108 |
| ## factor(ctylabel)Mauritius      | -4.713e+01 | 1.448e+01 | -3.254 |
| ## factor(ctylabel)Mexico         | -5.985e+01 | 1.895e+01 | -3.159 |
| ## factor(ctylabel)Moldova        | -4.977e+01 | 1.619e+01 | -3.075 |
| ## factor(ctylabel)Mongolia       | -4.933e+01 | 1.544e+01 | -3.194 |
| ## factor(ctylabel)Morocco        | -5.640e+01 | 1.787e+01 | -3.155 |
| ## factor(ctylabel)Mozambique     | -5.472e+01 | 1.742e+01 | -3.141 |
| ## factor(ctylabel)Nepal          | -5.712e+01 | 1.750e+01 | -3.264 |
| ## factor(ctylabel)Nicaragua      | -4.911e+01 | 1.590e+01 | -3.089 |
| ## factor(ctylabel)Niger          | -5.318e+01 | 1.663e+01 | -3.197 |
| ## factor(ctylabel)Nigeria        | -5.965e+01 | 1.928e+01 | -3.094 |
| ## factor(ctylabel)Oman           | -4.697e+01 | 1.536e+01 | -3.057 |
| ## factor(ctylabel)Pakistan       | -6.267e+01 | 1.936e+01 | -3.237 |
| ## factor(ctylabel)Panama         | -4.687e+01 | 1.536e+01 | -3.052 |
| ## factor(ctylabel)PapuaNewGuinea | -4.911e+01 | 1.587e+01 | -3.095 |
| ## factor(ctylabel)Paraguay       | -5.028e+01 | 1.589e+01 | -3.165 |
| ## factor(ctylabel)Peru           | -5.614e+01 | 1.753e+01 | -3.203 |
| ## factor(ctylabel)Philippines    | -6.012e+01 | 1.878e+01 | -3.201 |
| ## factor(ctylabel)Poland         | -5.772e+01 | 1.825e+01 | -3.163 |
| ## factor(ctylabel)Romania        | -5.658e+01 | 1.775e+01 | -3.187 |
| ## factor(ctylabel)Russia         | -6.337e+01 | 1.968e+01 | -3.220 |
| ## factor(ctylabel)Rwanda         | -5.250e+01 | 1.651e+01 | -3.180 |

|                                    |            |           |        |
|------------------------------------|------------|-----------|--------|
| ## factor(ctylabel)Senegal         | -5.220e+01 | 1.645e+01 | -3.174 |
| ## factor(ctylabel)SierraLeone     | -5.151e+01 | 1.604e+01 | -3.211 |
| ## factor(ctylabel)SouthAfrica     | -5.915e+01 | 1.841e+01 | -3.213 |
| ## factor(ctylabel)SriLanka        | -5.606e+01 | 1.735e+01 | -3.231 |
| ## factor(ctylabel)Swaziland       | -4.038e+01 | 1.440e+01 | -2.803 |
| ## factor(ctylabel)Syria           | -5.427e+01 | 1.712e+01 | -3.171 |
| ## factor(ctylabel)Tanzania        | -5.716e+01 | 1.809e+01 | -3.160 |
| ## factor(ctylabel)Thailand        | -5.950e+01 | 1.853e+01 | -3.212 |
| ## factor(ctylabel)Togo            | -4.874e+01 | 1.579e+01 | -3.087 |
| ## factor(ctylabel)Trinidad&Tobago | -4.314e+01 | 1.453e+01 | -2.970 |
| ## factor(ctylabel)Tunisia         | -5.122e+01 | 1.668e+01 | -3.070 |
| ## factor(ctylabel)Turkey          | -6.083e+01 | 1.852e+01 | -3.284 |
| ## factor(ctylabel)Uganda          | -5.539e+01 | 1.749e+01 | -3.166 |
| ## factor(ctylabel)Ukraine         | -6.001e+01 | 1.870e+01 | -3.209 |
| ## factor(ctylabel)Uruguay         | -5.040e+01 | 1.557e+01 | -3.236 |
| ## factor(ctylabel)Venezuela       | -5.669e+01 | 1.735e+01 | -3.268 |
| ## factor(ctylabel)Zambia          | -5.067e+01 | 1.670e+01 | -3.035 |
| ## factor(ctylabel)Zimbabwe        | -5.362e+01 | 1.681e+01 | -3.190 |
| ##                                 | Pr(> t )   |           |        |
| ## l1polity                        | 0.214252   |           |        |
| ## l1polity_squared                | 0.255484   |           |        |
| ## l1signed                        | 0.446403   |           |        |
| ## l1office                        | 0.219355   |           |        |
| ## l1gdp_pc                        | 0.636568   |           |        |
| ## l1lnpop                         | 0.000901   | ***       |        |
| ## l1ecris2                        | 0.856053   |           |        |
| ## l1bpc1                          | 0.221207   |           |        |
| ## l1avnewtar                      | 0.000412   | ***       |        |
| ## factor(ctylabel)Albania         | 0.001850   | **        |        |
| ## factor(ctylabel)Algeria         | 0.001259   | **        |        |
| ## factor(ctylabel)Angola          | 0.290862   |           |        |
| ## factor(ctylabel)Argentina       | 0.001386   | **        |        |
| ## factor(ctylabel)Armenia         | 0.002692   | **        |        |
| ## factor(ctylabel)Azerbaijan      | 0.019074   | *         |        |

|   |             |
|---|-------------|
| ## factor(ctylabel)Bangladesh             | 0.001181 ** |
| ## factor(ctylabel)Belarus                | 0.001426 ** |
| ## factor(ctylabel)Benin                  | 0.001399 ** |
| ## factor(ctylabel)Bhutan                 | 0.001532 ** |
| ## factor(ctylabel)Bolivia                | 0.001744 ** |
| ## factor(ctylabel)Botswana               | 0.002086 ** |
| ## factor(ctylabel)Brazil                 | 0.001252 ** |
| ## factor(ctylabel)Bulgaria               | 0.001570 ** |
| ## factor(ctylabel)Burundi                | 0.001388 ** |
| ## factor(ctylabel)Cambodia               | 0.004563 ** |
| ## factor(ctylabel)Cameroon               | 0.001652 ** |
| ## factor(ctylabel)CentralAfricanRepublic | 0.001467 ** |
| ## factor(ctylabel)Chad                   | 0.001672 ** |
| ## factor(ctylabel)Chile                  | 0.001836 ** |
| ## factor(ctylabel)China                  | 0.001778 ** |
| ## factor(ctylabel)Colombia               | 0.001221 ** |
| ## factor(ctylabel)Comoros                | 0.001587 ** |
| ## factor(ctylabel)Congo                  | 0.001719 ** |
| ## factor(ctylabel)CostaRica              | 0.001596 ** |
| ## factor(ctylabel)Coted'Ivoire           | 0.001852 ** |
| ## factor(ctylabel)Croatia                | 0.002292 ** |
| ## factor(ctylabel)DominicanRepublic      | 0.001590 ** |
| ## factor(ctylabel)Ecuador                | 0.001327 ** |
| ## factor(ctylabel)Egypt                  | 0.001884 ** |
| ## factor(ctylabel)ElSalvador             | 0.001229 ** |
| ## factor(ctylabel)EquatorialGuinea       | 0.431734    |
| ## factor(ctylabel)Estonia                | 0.004542 ** |
| ## factor(ctylabel)Ethiopia               | 0.001424 ** |
| ## factor(ctylabel)Fiji                   | 0.002305 ** |
| ## factor(ctylabel)Gabon                  | 0.002639 ** |
| ## factor(ctylabel)Gambia                 | 0.001880 ** |
| ## factor(ctylabel)Georgia                | 0.007015 ** |
| ## factor(ctylabel)Ghana                  | 0.001424 ** |
| ## factor(ctylabel)Guatemala              | 0.001472 ** |

|                                   |          |     |
|-----------------------------------|----------|-----|
| ## factor(ctylabel)Guinea         | 0.001529 | **  |
| ## factor(ctylabel)GuineaBissau   | 0.001733 | **  |
| ## factor(ctylabel)Guyana         | 0.005025 | **  |
| ## factor(ctylabel)Haiti          | 0.001482 | **  |
| ## factor(ctylabel)Honduras       | 0.001452 | **  |
| ## factor(ctylabel)Hungary        | 0.002230 | **  |
| ## factor(ctylabel)India          | 0.000942 | *** |
| ## factor(ctylabel)Indonesia      | 0.001592 | **  |
| ## factor(ctylabel)Iran           | 0.001371 | **  |
| ## factor(ctylabel)Jamaica        | 0.001477 | **  |
| ## factor(ctylabel)Jordan         | 0.001960 | **  |
| ## factor(ctylabel)Kazakhstan     | 0.003045 | **  |
| ## factor(ctylabel)Kenya          | 0.001497 | **  |
| ## factor(ctylabel)Korea          | 0.001363 | **  |
| ## factor(ctylabel)KyrgyzRepublic | 0.003441 | **  |
| ## factor(ctylabel)Laos           | 0.002910 | **  |
| ## factor(ctylabel)Latvia         | 0.003821 | **  |
| ## factor(ctylabel)Lesotho        | 0.003739 | **  |
| ## factor(ctylabel)Lithuania      | 0.002126 | **  |
| ## factor(ctylabel)Madagascar     | 0.001347 | **  |
| ## factor(ctylabel)Malawi         | 0.001628 | **  |
| ## factor(ctylabel)Malaysia       | 0.002482 | **  |
| ## factor(ctylabel)Mali           | 0.001476 | **  |
| ## factor(ctylabel)Mauritania     | 0.001917 | **  |
| ## factor(ctylabel)Mauritius      | 0.001163 | **  |
| ## factor(ctylabel)Mexico         | 0.001616 | **  |
| ## factor(ctylabel)Moldova        | 0.002143 | **  |
| ## factor(ctylabel)Mongolia       | 0.001429 | **  |
| ## factor(ctylabel)Morocco        | 0.001633 | **  |
| ## factor(ctylabel)Mozambique     | 0.001715 | **  |
| ## factor(ctylabel)Nepal          | 0.001123 | **  |
| ## factor(ctylabel)Nicaragua      | 0.002045 | **  |
| ## factor(ctylabel)Niger          | 0.001416 | **  |
| ## factor(ctylabel)Nigeria        | 0.002013 | **  |



```

## factor(ctylabel)Oman                0.002270 **
## factor(ctylabel)Pakistan            0.001235 **
## factor(ctylabel)Panama              0.002313 **
## factor(ctylabel)PapuaNewGuinea      0.002002 **
## factor(ctylabel)Paraguay            0.001580 **
## factor(ctylabel)Peru                0.001388 **
## factor(ctylabel)Philippines         0.001396 **
## factor(ctylabel)Poland              0.001593 **
## factor(ctylabel)Romania             0.001466 **
## factor(ctylabel)Russia              0.001309 **
## factor(ctylabel)Rwanda              0.001502 **
## factor(ctylabel)Senegal             0.001535 **
## factor(ctylabel)SierraLeone         0.001350 **
## factor(ctylabel)SouthAfrica         0.001342 **
## factor(ctylabel)SriLanka           0.001258 **
## factor(ctylabel)Swaziland           0.005119 **
## factor(ctylabel)Syria               0.001551 **
## factor(ctylabel)Tanzania            0.001608 **
## factor(ctylabel)Thailand            0.001346 **
## factor(ctylabel)Togo               0.002055 **
## factor(ctylabel)Trinidad&Tobago     0.003024 **
## factor(ctylabel)Tunisia            0.002179 **
## factor(ctylabel)Turkey             0.001045 **
## factor(ctylabel)Uganda              0.001574 **
## factor(ctylabel)Ukraine            0.001358 **
## factor(ctylabel)Uruguay            0.001237 **
## factor(ctylabel)Venezuela          0.001108 **
## factor(ctylabel)Zambia              0.002449 **
## factor(ctylabel)Zimbabwe           0.001452 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.963 on 1546 degrees of freedom
## (3712 observations deleted due to missingness)

```

```
## Multiple R-squared:  0.3616, Adjusted R-squared:  0.3153
## F-statistic: 7.817 on 112 and 1546 DF,  p-value: < 2.2e-16

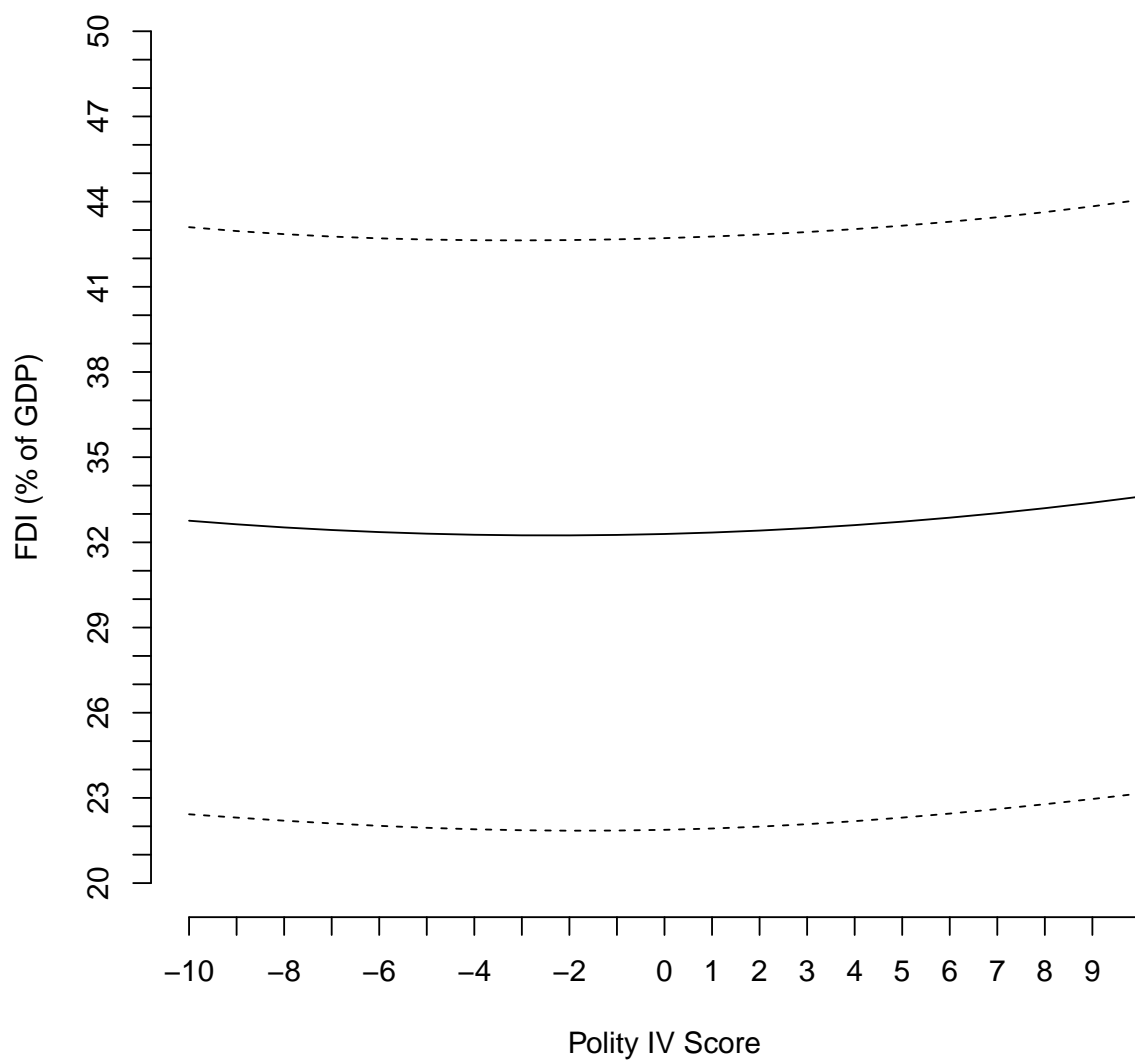
nd <- data.frame(l1polity = seq(-10, 10, by = 1), l1polity_squared = seq(-10,
  10, by = 1)^2, l1signed = rep(0.1511, 21), l1office = rep(8.431, 21), l1gdp_pc = rep
  21), l1lnpop = rep(15.1, 21), l1ecris2 = rep(0.0641, 21), l1bpc1 = rep(0.5909,
  21), l1avnewtar = rep(14.91, 21), ctylabel = rep("Angola", 21))

pred.p1 <- predict(lm_fdi2, type = "response", se.fit = TRUE, newdata = nd)
pred.table <- cbind(pred.p1$fit, pred.p1$se.fit)

fit <- pred.p1$fit
low <- pred.p1$fit - 2 * pred.p1$se.fit
high <- pred.p1$fit + 2 * pred.p1$se.fit
cis <- cbind(fit, low, high)

plot(pred.p1$fit, type = "l", ylim = c(20, 50), main = "Polity IV Score and FDI (% GDP)
  xlab = "Polity IV Score", ylab = "FDI (% of GDP)", axes = FALSE)
axis(1, at = seq(1, 21), labels = seq(-10, 10, 1))
axis(2, at = seq(20, 50), labels = seq(20, 50))
matlines(cis[, c(2, 3)], lty = 2, col = "black")
```

### Polity IV Score and FDI (% GDP) (Angola)



## Problem 2

```
### a
setwd("C:/Users/Jan/OneDrive/Documents/GitHub/ps630_lab/")
library(readstata13)
na_data = read.dta13("na_data.dta")
summary(na_data)
```

```

## countrycode          year          v_c
## Length:10624      Min.    :1950      Min.    :0.000e+00
## Class :character  1st Qu.:1973      1st Qu.:7.600e+02
## Mode  :character  Median :1986      Median :1.296e+04
##                               Mean  :1985      Mean  :8.369e+06
##                               3rd Qu.:1999      3rd Qu.:2.230e+05
##                               Max.   :2011      Max.   :4.053e+09
##                               NA's    :561
##
##      v_i              v_g              v_x
## Min.   :      -7427  Min.   :         0  Min.   :0.000e+00
## 1st Qu.:       251  1st Qu.:       171  1st Qu.:4.260e+02
## Median :      4319  Median :      3197  Median :5.520e+03
## Mean   :    4418919  Mean   :   1658085  Mean   :4.787e+06
## 3rd Qu.:     70963  3rd Qu.:   49433  3rd Qu.:8.595e+04
## Max.   :2433863510  Max.   :667440135  Max.   :1.955e+09
## NA's   :561        NA's   :561        NA's   :519
##
##      v_m              v_gdp              q_c
## Min.   :0.000e+00  Min.   :0.000e+00  Min.   :2.000e+00
## 1st Qu.:5.320e+02  1st Qu.:1.188e+03  1st Qu.:7.238e+03
## Median :6.670e+03  Median :2.077e+04  Median :1.263e+05
## Mean   :4.538e+06  Mean   :1.473e+07  Mean   :1.202e+07
## 3rd Qu.:1.094e+05  3rd Qu.:3.398e+05  3rd Qu.:9.104e+05
## Max.   :2.194e+09  Max.   :7.427e+09  Max.   :2.343e+09
## NA's   :519        NA's   :519        NA's   :561
##
##      q_i              q_g              q_x
## Min.   :   -39042  Min.   :         6  Min.   :0.000e+00
## 1st Qu.:    2543  1st Qu.:    1865  1st Qu.:3.730e+03
## Median :    39263  Median :    27296  Median :4.116e+04
## Mean   :   6616470  Mean   :   2764949  Mean   :7.191e+06
## 3rd Qu.:   279383  3rd Qu.:   224138  3rd Qu.:2.672e+05
## Max.   :1098261440  Max.   :372916568  Max.   :1.453e+09
## NA's   :561        NA's   :561        NA's   :519
##
##      q_m              q_gdp              pop
## Min.   :7.000e+00  Min.   :9.000e+00  Min.   :4.608e+03

```

```
## 1st Qu.:5.002e+03 1st Qu.:1.252e+04 1st Qu.:7.673e+05
## Median :4.708e+04 Median :2.065e+05 Median :4.951e+06
## Mean :6.188e+06 Mean :2.188e+07 Mean :3.177e+07
## 3rd Qu.:3.579e+05 3rd Qu.:1.424e+06 3rd Qu.:1.614e+07
## Max. :1.223e+09 Max. :3.903e+09 Max. :1.324e+09
## NA's :519 NA's :519 NA's :459
## xr xr2 v_gfcf
## Min. : 0.00 Min. : 0.00 Min. :0.000e+00
## 1st Qu.: 0.90 1st Qu.: 0.91 1st Qu.:3.900e+02
## Median : 2.57 Median : 2.64 Median :7.822e+03
## Mean : 220.40 Mean : 221.12 Mean :4.854e+06
## 3rd Qu.: 28.58 3rd Qu.: 31.64 3rd Qu.:1.101e+05
## Max. :31900.00 Max. :31900.00 Max. :2.378e+09
## NA's :459 NA's :459 NA's :2370
## q_gfcf
## Min. :3.000e+00
## 1st Qu.:1.826e+03
## Median :4.209e+04
## Mean :6.118e+06
## 3rd Qu.:3.044e+05
## Max. :1.004e+09
## NA's :2390

### b
na_data$gdpgrowth = NA
for (i in 2:length(na_data$q_gdp)) {
  if (na_data$countrycode[i] == na_data$countrycode[i - 1]) {
    na_data$gdpgrowth[i] = (na_data$q_gdp[i]/na_data$q_gdp[i - 1] - 1) *
      100
  }
}

summary(na_data$gdpgrowth)

## Min. 1st Qu. Median Mean 3rd Qu. Max. NA's
```

```

## -66.120    1.372    4.038    3.997    6.776 205.000      728

#### c
na_data$date = na_data$year

LDC$countrycode = NA
LDC$countrycode[LDC$ctylabel == "Turkey"] = "TUR"
LDC$countrycode[LDC$ctylabel == "SouthAfrica"] = "ZAF"
LDC$countrycode[LDC$ctylabel == "Mexico"] = "MEX"

merged_data = merge(LDC, na_data, by = c("countrycode", "date"))

newmodel = lm(newtar ~ l1polity + gdpgrowth + factor(countrycode) - 1, data = merged_data)
summary(newmodel)

##
## Call:
## lm(formula = newtar ~ l1polity + gdpgrowth + factor(countrycode) -
##      1, data = merged_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -9.4339 -3.7056 -0.2797  4.1416 11.7623
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## l1polity          -1.4270     0.2978  -4.792 3.40e-05 ***
## gdpgrowth           0.4709     0.2907   1.620  0.115
## factor(countrycode)MEX 15.1326     1.5137   9.997 1.63e-11 ***
## factor(countrycode)TUR 29.2080     2.7362  10.675 3.08e-12 ***
## factor(countrycode)ZAF 20.5447     2.5459   8.070 2.59e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.694 on 33 degrees of freedom

```

```
## (52 observations deleted due to missingness)
## Multiple R-squared: 0.9176, Adjusted R-squared: 0.9051
## F-statistic: 73.52 on 5 and 33 DF, p-value: < 2.2e-16
```

### Problem 3

a) As grader, please make sure that the person who has submitted the homework has answered all questions of this problem. This includes a brief explanation (2-3 sentences) of the student's theory and a reference to both the dependent and key independent variable.

Make sure that the students has done the following things:

1. The student has explained which datasets contain the variables and provided an overview of how the variables are coded there.
2. If and only if there were variables that were nominal or ordinal scale or coded as characters, the student has recognized that these variables have to be transformed to be used for a linear regression. Nominal variables have to be introduced as factors (dummies). Ordinal variables have to be either used as factors (dummies) or have to be assigned numerical values. Similarly, variables that are coded as characters have to be either introduced as factors (dummies) or recoded as numerical variables.
3. The students has discussed the numbers of units and the time periods covered by the datasets.

b) As grader, please make sure that the person who has submitted the homework has answered all questions of this problem. This includes a brief theoretical explanation for the importance of at least two control variables that the student suggests to use for the final paper. For all control variables there should be a brief reference to either literature that has used the variable, the concept of omitted variable bias, or both.

Make sure that the students has done the following things:

1. The student has explained which datasets contain the variables and provided an overview of how the variables are coded there.
2. If and only if there were variables that were nominal or ordinal scale or coded as characters, the student has recognized that these variables have to be transformed to

be used for a linear regression. Nominal variables have to be introduced as factors (dummies). Ordinal variables have to be either used as factors (dummies) or have to be assigned numerical values. Similarly, variables that are coded as characters have to be either introduced as factors (dummies) or recoded as numerical variables.

3. The students has addressed potential differences in the time periods and units covered. For example, data for the control variables may be available only for OECD countries while the data for the dependent variable may only be available for developing countries. Another example would be that data for the control variables may be available on a quarterly basis while data for the dependent variable may be available on an annual basis.
4. The students has addressed differences in the coding of time periods and units. For example, the names of countries are coded as full names in one dataset while another dataset uses 3-letter isocodes to refer to countries. Another example would be that time in one dataset could be coded in the format YYYY-MM (Y = year, M = month) while it could be coded in the format YY-MM in another dataset.

## Statistical Theory: Omitted Variable Bias

### Problem 4

a)

b)