R Notebook

```
load("nunn.Rda")
```

ln_realgdp2000 => outcome, Y

atlantic_dist => IV original, Z1 indian_dist => IV original, Z2 saharan_dist => IV original, Z3 redsea_dist => IV original, Z4

ln_export_ area => treatment original, D

low_distance => IV from paper, Z high_slavery => treatment paper, D

a)

i. coutries that are complier is when Z = 0, D = 0, and Z =1, D =1 countries that are always takers is regardless if Z = 1 or 0, D = 1
ii. Calculate and report the proportion of compliers and the intent-to-treat effect. itt is -0.0159, prop.c is 0.0578 => 5.78% are compliers
iii. find the CACE CACE is -0.2757
iv. it is not statistically significant

```
prop.c <- sum(nunn$high_slavery[nunn$low_distance==1])/length(nunn$high_slavery[nunn$low_distance==1]) -

itt <- mean(nunn$ln_realgdp2000[nunn$low_distance==1]) - mean(nunn$ln_realgdp2000[nunn$low_distance==0])

cace <- itt/prop.c

library('AER')
```

```
## Loading required package: car
## Loading required package: carData
## Loading required package: lmtest
## Loading required package: zoo
##
## Attaching package: 'zoo'
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
## Loading required package: sandwich
## Loading required package: survival
```

```
summary(ivreg(ln_realgdp2000 ~ high_slavery | low_distance, data = nunn))
```

```
##
## Call:
## ivreg(formula = ln_realgdp2000 ~ high_slavery | low_distance,
##     data = nunn)
##
```

```
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.8180 -0.5012 -0.1913  0.5120  2.0710
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    7.2025     0.9762   7.378 1.54e-09 ***
## high_slavery  -0.2757     3.8791  -0.071    0.944
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.8075 on 50 degrees of freedom
## Multiple R-Squared: 0.06148, Adjusted R-squared: 0.04271
## Wald test: 0.005051 on 1 and 50 DF,  p-value: 0.9436
```

b) Turning now to the analysis that Nunn conducts in his paper, replicate the first-stage results from the first column of Table IV on p.162. Report your results.

```
summary(lm(ln_export_area ~ atlantic_dist + indian_dist + saharan_dist + redsea_dist, data = nunn))
```

```
##
## Call:
## lm(formula = ln_export_area ~ atlantic_dist + indian_dist + saharan_dist +
##     redsea_dist, data = nunn)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.3574 -2.4772  0.2513  2.8323  5.9544
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)   29.10971    6.95941   4.183 0.000125 ***
## atlantic_dist -1.31399    0.35678  -3.683 0.000594 ***
## indian_dist   -1.09544    0.37978  -2.884 0.005901 **
## saharan_dist  -2.43487    0.82305  -2.958 0.004830 **
## redsea_dist   -0.00186    0.71041  -0.003 0.997922
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.445 on 47 degrees of freedom
## Multiple R-squared:  0.2789, Adjusted R-squared:  0.2176
## F-statistic: 4.545 on 4 and 47 DF,  p-value: 0.003472
```

c) Are the instruments in this paper subject to the weak instrument problem? What consequences does this have, if any, for our interpretation of the results? Explain your answer, providing evidence from the data.

```
library('lmtest')
mod1 <- lm(ln_export_area ~ factor(colonial_power) + equator_dist + longitude + rain_min + humid_max + l
mod2 <- lm(ln_export_area ~ atlantic_dist + indian_dist + saharan_dist + redsea_dist + factor(colonial_p
waldtest(mod2, mod1)
```

```
## Wald test
##
## Model 1: ln_export_area ~ atlantic_dist + indian_dist + saharan_dist +
##     redsea_dist + factor(colonial_power) + equator_dist + longitude +
```

```
##     rain_min + humid_max + low_temp + ln_coastline_area
## Model 2: ln_export_area ~ factor(colonial_power) + equator_dist + longitude +
##     rain_min + humid_max + low_temp + ln_coastline_area
##   Res.Df Df      F Pr(>F)
## 1     34
## 2     38 -4 0.8111 0.5269
```

d) Do you think that the instruments in this example satisfy the exclusion restriction assumption? Briefly explain your answer.

```
summary(lm(ln_realgdp2000 ~  atlantic_dist + indian_dist + saharan_dist + redsea_dist, data = nunn))
```

```
##
## Call:
## lm(formula = ln_realgdp2000 ~ atlantic_dist + indian_dist + saharan_dist +
##     redsea_dist, data = nunn)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.48278 -0.47593  0.00978  0.32312  1.96445
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.14859    1.39944   0.106 0.915893
## atlantic_dist  0.32843    0.07174   4.578 3.45e-05 ***
## indian_dist    0.30710    0.07637   4.021 0.000208 ***
## saharan_dist   0.59358    0.16550   3.586 0.000795 ***
## redsea_dist    0.10075    0.14285   0.705 0.484105
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6928 on 47 degrees of freedom
## Multiple R-squared:  0.3507, Adjusted R-squared:  0.2954
## F-statistic: 6.345 on 4 and 47 DF,  p-value: 0.0003621
```

e) Replicate the second-stage coefficients and standard errors for ln(exports/area) in columns (1), (2) and (3) of Table IV on p.162 of the paper. Report your results and briefly interpret each of the three estimated LATEs.

```
#model 1
summary(ivreg(ln_realgdp2000 ~ ln_export_area | atlantic_dist + indian_dist + saharan_dist + redsea_dis
```

```
##
## Call:
## ivreg(formula = ln_realgdp2000 ~ ln_export_area | atlantic_dist +
##     indian_dist + saharan_dist + redsea_dist, data = nunn)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.9254 -0.4602  0.1429  0.4917  1.4163
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)     7.81135    0.20375  38.337  < 2e-16 ***
## ln_export_area -0.20794    0.05301  -3.923 0.000267 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7787 on 50 degrees of freedom
## Multiple R-Squared: 0.1273,  Adjusted R-squared: 0.1098
## Wald test: 15.39 on 1 and 50 DF,  p-value: 0.0002674
```

```
#model 2
summary(ivreg(ln_realgdp2000 ~ ln_export_area + colonial_power | atlantic_dist + indian_dist + saharan_
```

```
##
## Call:
## ivreg(formula = ln_realgdp2000 ~ ln_export_area + colonial_power |
##     atlantic_dist + indian_dist + saharan_dist + redsea_dist +
##         colonial_power, data = nunn)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.95861 -0.44487  0.07661  0.43823  1.33770
##
## Coefficients:
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)               6.4640     0.4271  15.136  < 2e-16 ***
## ln_export_area           -0.2014     0.0472  -4.267 0.000107 ***
## colonial_powerFrance      1.3957     0.4612   3.026 0.004174 **
## colonial_powerGermany     1.4824     0.8535   1.737 0.089586 .
## colonial_powerItaly       1.6115     0.8416   1.915 0.062204 .
## colonial_powernone        1.2520     0.6923   1.808 0.077547 .
## colonial_powerPortugal    1.2939     0.5420   2.387 0.021447 *
## colonial_powerSpain       2.3195     0.8500   2.729 0.009174 **
## colonial_powerUK          1.4033     0.4615   3.040 0.004013 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7289 on 43 degrees of freedom
## Multiple R-Squared: 0.3424,  Adjusted R-squared: 0.2201
## Wald test: 4.322 on 8 and 43 DF,  p-value: 0.0006999
```

```
#model 3
summary(ivreg(ln_realgdp2000 ~ ln_export_area + colonial_power + equator_dist + longitude + rain_min + h
```

```
##
## Call:
## ivreg(formula = ln_realgdp2000 ~ ln_export_area + colonial_power +
##     equator_dist + longitude + rain_min + humid_max + low_temp +
##     ln_coastline_area | atlantic_dist + indian_dist + saharan_dist +
##     redsea_dist + colonial_power + equator_dist + longitude +
##     rain_min + humid_max + low_temp + ln_coastline_area, data = nunn)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.77651 -0.45102  0.05974  0.54425  1.75322
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)              6.978903   1.816685   3.842 0.000463 ***
## ln_export_area          -0.285659   0.152702  -1.871 0.069312 .
```

```
## colonial_powerFrance    1.487015   0.910976   1.632 0.111093
## colonial_powerGermany   1.343497   1.412438   0.951 0.347680
## colonial_powerItaly     1.792295   1.451631   1.235 0.224735
## colonial_powernone      1.622853   1.683846   0.964 0.341415
## colonial_powerPortugal  1.101675   1.025801   1.074 0.289792
## colonial_powerSpain     1.712097   1.238201   1.383 0.175037
## colonial_powerUK        1.682392   0.985964   1.706 0.096326 .
## equator_dist           -0.007522   0.030568  -0.246 0.806999
## longitude              -0.010117   0.011274  -0.897 0.375335
## rain_min               -0.011930   0.013009  -0.917 0.365082
## humid_max              -0.002182   0.019982  -0.109 0.913634
## low_temp                0.020548   0.051820   0.397 0.693987
## ln_coastline_area       0.023054   0.073814   0.312 0.756551
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9488 on 37 degrees of freedom
## Multiple R-Squared: 0.04113, Adjusted R-squared: -0.3217
## Wald test: 1.703 on 14 and 37 DF,  p-value: 0.09745
```