

# POLS/CS&SS 503: Assignment 3

*Jeffrey Arnold*

*May 17, 2017*

## Setup

```
library("tidyverse")

## Loading tidyverse: ggplot2
## Loading tidyverse: tibble
## Loading tidyverse: tidyr
## Loading tidyverse: readr
## Loading tidyverse: purrr
## Loading tidyverse: dplyr

## Conflicts with tidy packages -----
## filter(): dplyr, stats
## lag():    dplyr, stats
library("haven")
library("AER")

## Loading required package: car
##
## Attaching package: 'car'
## The following object is masked from 'package:dplyr':
##
##     recode
## The following object is masked from 'package:purrr':
##
##     some
## 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
library("Formula")
```

## 0.1 Problem 1

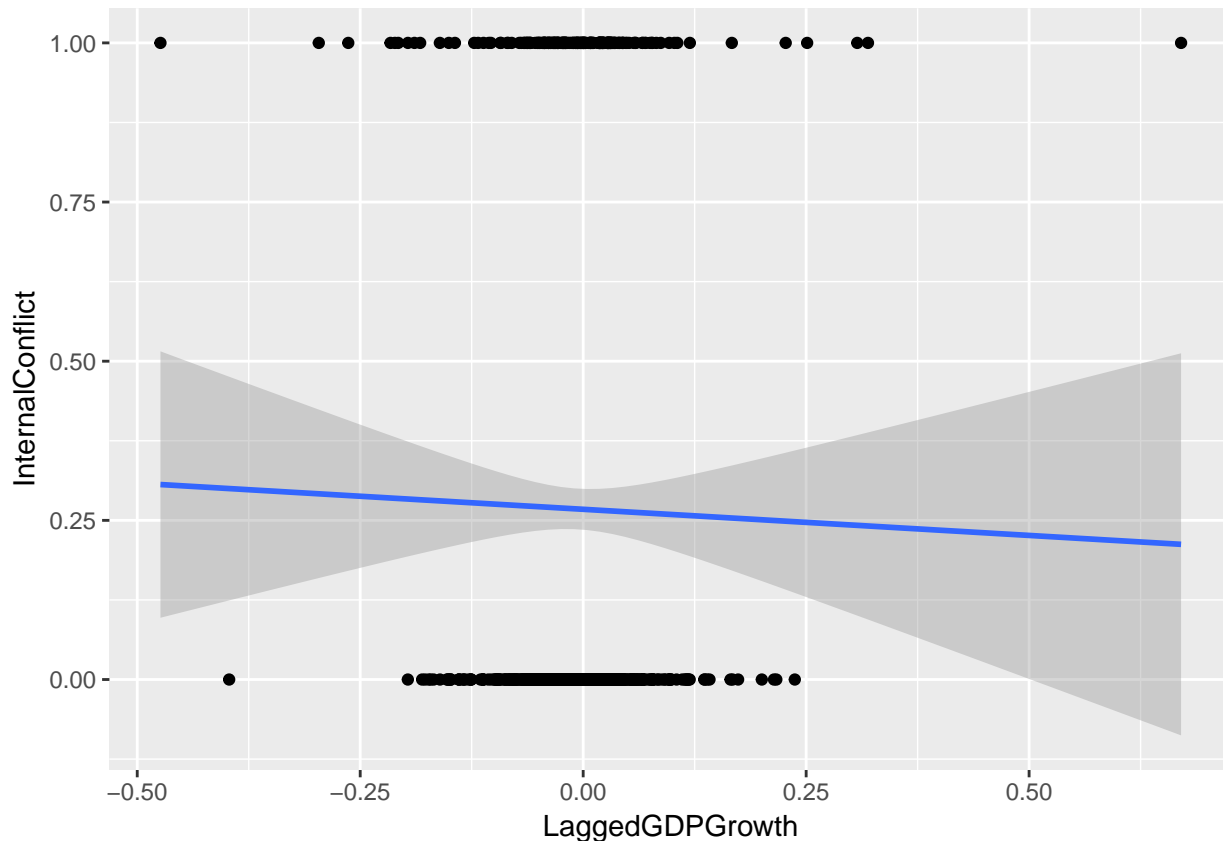
Bailey (2016) Ex 9.1

```
RainIV <- read_csv("Ch09.Ex1.CivilWarRainInstrumentData/RainIV.csv")
```

```
## Parsed with column specification:
## cols(
##   .default = col_integer(),
##   country_name = col_character(),
##   country_code = col_character(),
##   GPCP = col_double(),
##   RainfallGrowth = col_double(),
##   LaggedRainfallGrowth = col_double(),
##   pop = col_double(),
##   lpopl1 = col_double(),
##   Mountains = col_double(),
##   lmtnest = col_double(),
##   EthnicFrac = col_double(),
##   ReligiousFrac = col_double(),
##   GDPGrowth = col_double(),
##   LaggedGDPGrowth = col_double(),
##   InitialGDP = col_double()
## )
```

```
## See spec(...) for full column specifications.
```

```
ggplot(RainIV, aes(y = InternalConflict, x = LaggedGDPGrowth)) +
  geom_point() +
  geom_smooth(method = "lm")
```



```
mod1_f <- InternalConflict ~ LaggedGDPGrowth
mod1 <- lm(mod1_f, data = RainIV)
summary(mod1)
```

```
##
## Call:
## lm(formula = mod1_f, data = RainIV)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.2999 -0.2689 -0.2660  0.7228  0.7876
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.26738    0.01631  16.389  <2e-16 ***
## LaggedGDPGrowth -0.08206    0.22485  -0.365    0.715
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4434 on 741 degrees of freedom
## Multiple R-squared:  0.0001797, Adjusted R-squared: -0.00117
## F-statistic: 0.1332 on 1 and 741 DF, p-value: 0.7152
```

b. Add controls for initial GDP (InitialGDP), democracy (‘) mountains, and ethnic and religious fractionalization to the model.

```
mod2_f <- update(mod1_f, . ~ . + InitialGDP + Democracy +
                  Mountains + EthnicFrac + ReligiousFrac)
```

```
mod2 <- lm(mod2_f, data = RainIV)
summary(mod2)
```

```
##
## Call:
## lm(formula = mod2_f, data = RainIV)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.5654 -0.2811 -0.2221  0.4570  0.9459
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.0703555  0.0731012   0.962  0.33614
## LaggedGDPGrowth -0.1087977  0.2200999  -0.494  0.62123
## InitialGDP      -0.0569091  0.0182258  -3.122  0.00186 **
## Democracy        0.0012242  0.0028894   0.424  0.67193
## Mountains        0.0038654  0.0009527   4.057 5.49e-05 ***
## EthnicFrac       0.3247931  0.0918181   3.537 0.00043 ***
## ReligiousFrac    0.0105162  0.0958907   0.110  0.91270
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4334 on 736 degrees of freedom
## Multiple R-squared:  0.05106,    Adjusted R-squared:  0.04332
## F-statistic:  6.6 on 6 and 736 DF,  p-value: 8.276e-07
```

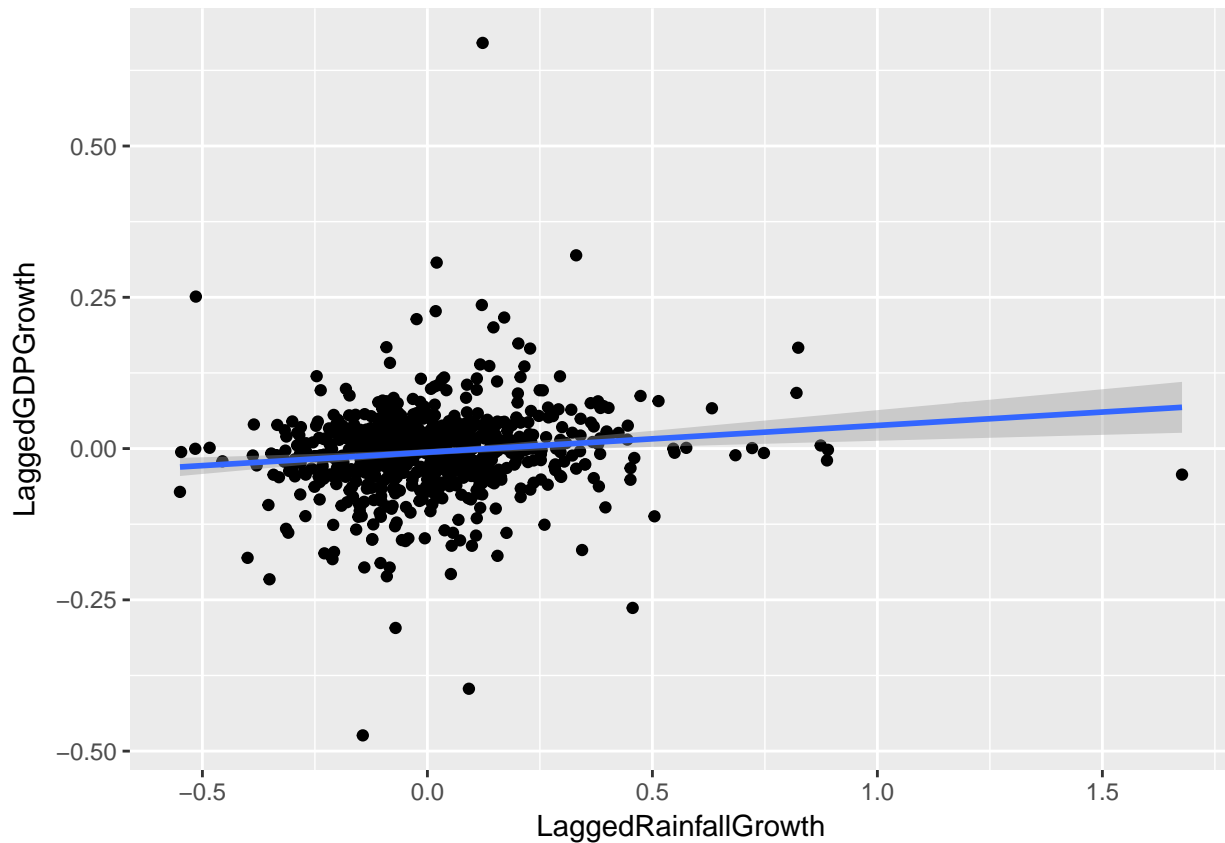
No. The coefficient on  $\beta$  would only be a causal estimate under an assumption of no omitted confounders. This seems highly unlikely.

c. The two conditions needed for a good instrument are

- inclusion restriction: Rainfall must be correlated with economic growth
- exclusion restriction: Rainfall must not be correlated with war except through its effect on economic growth.

The inclusion restriction can be tested (in the first stage):

```
ggplot(RainIV, aes(x = LaggedRainfallGrowth, y = LaggedGDPGrowth)) +
  geom_point() +
  geom_smooth(method = "lm")
```



The exclusion restriction cannot be tested statistically, and must be argued outside the model.

d. Instrumenting GDP growth with rainfall randomly assigns some part of GDP growth to countries.

e.

```
mod3_f <- InternalConflict ~ LaggedGDPGrowth + InitialGDP + Democracy + Mountains + EthnicFrac + ReligiousFrac
mod3 <- ivreg(mod3_f, data = RainIV)
summary(mod3)
```

```
##
## Call:
## ivreg(formula = mod3_f, data = RainIV)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.1693 -0.3106 -0.1897  0.4203  2.0093
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.062506   0.077268   0.809 0.418802
## LaggedGDPGrowth -2.063153   1.845106  -1.118 0.263857
## InitialGDP    -0.058080   0.019209  -3.024 0.002584 **
## Democracy      0.002361   0.003221   0.733 0.463785
## Mountains      0.004069   0.001020   3.988 7.34e-05 ***
## EthnicFrac     0.328851   0.096686   3.401 0.000707 ***
## ReligiousFrac  0.004724   0.101042   0.047 0.962721
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 0.456 on 736 degrees of freedom
## Multiple R-Squared: -0.05059, Adjusted R-squared: -0.05916
## Wald test: 6.133 on 6 and 736 DF, p-value: 2.748e-06
```

The coefficient on LaggedGDPGrowth is 20 times larger than in the OLS regression. It is still not statistically significant.

f. Redo the 2SLS with country fixed effects.

```
mod4_f <- InternalConflict ~ LaggedGDPGrowth + InitialGDP + Democracy + Mountains + EthnicFrac + ReligiousFrac
mod4 <- ivreg(mod4_f, data = RainIV)
summary(mod4)
```

```
##
## Call:
## ivreg(formula = mod4_f, data = RainIV)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.60872 -0.18282 -0.01501  0.13649  1.92662
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -0.752079   0.473516  -1.588  0.11267
## LaggedGDPGrowth -2.853380   1.535631  -1.858  0.06357 .
## InitialGDP    -0.011809   0.072129  -0.164  0.86999
## Democracy       0.001065   0.003276   0.325  0.74518
## Mountains      0.098156   0.017018   5.768 1.21e-08 ***
## EthnicFrac     0.412905   0.376584   1.096  0.27326
## ReligiousFrac  0.749920   0.871221   0.861  0.38966
## country_codeBDI -6.579658   1.236583  -5.321 1.39e-07 ***
## country_codeBEN  0.166612   0.116083   1.435  0.15165
## country_codeBFA  0.234677   0.134340   1.747  0.08110 .
## country_codeBWA  0.251482   0.166556   1.510  0.13152
## country_codeCAF -0.580915   0.217323  -2.673  0.00769 **
## country_codeCIV -0.114346   0.165163  -0.692  0.48896
## country_codeCMR -1.768851   0.200940  -8.803 < 2e-16 ***
## country_codeCOG  0.289731   0.144290   2.008  0.04503 *
## country_codeDJI  0.101831   0.358274   0.284  0.77632
## country_codeETH -6.086221   1.112293  -5.472 6.21e-08 ***
## country_codeGAB  0.162233   0.283837   0.572  0.56780
## country_codeGHA  0.073439   0.223433   0.329  0.74249
## country_codeGIN  0.064057   0.263453   0.243  0.80796
## country_codeGMB  0.334810   0.288195   1.162  0.24573
## country_codeGNB  0.123575   0.122652   1.008  0.31403
## country_codeKEN -2.632596   0.305915  -8.606 < 2e-16 ***
## country_codeLBR  0.104590   0.173242   0.604  0.54622
## country_codeLSO -7.567227   1.466494  -5.160 3.22e-07 ***
## country_codeMDG -2.976911   0.558717  -5.328 1.34e-07 ***
## country_codeMLI  0.368259   0.321301   1.146  0.25213
## country_codeMOZ  0.423353   0.145244   2.915  0.00367 **
## country_codeMRT  0.632220   0.418388   1.511  0.13122
## country_codeMWI -0.812187   0.124855  -6.505 1.48e-10 ***
## country_codeNAM -0.548080   0.272637  -2.010  0.04478 *
## country_codeNER  0.266671   0.216693   1.231  0.21887
```

```
## country_codeNGA -0.310737 0.118896 -2.614 0.00915 **
## country_codeRWA -6.395050 1.222877 -5.230 2.25e-07 ***
## country_codeSDN 0.416879 0.158764 2.626 0.00883 **
## country_codeSEN 0.726010 0.301963 2.404 0.01646 *
## country_codeSLE 0.239721 0.107686 2.226 0.02632 *
## country_codeSOM 0.447733 0.598064 0.749 0.45433
## country_codeSWZ -1.232084 0.248451 -4.959 8.90e-07 ***
## country_codeTCD -0.034781 0.111964 -0.311 0.75616
## country_codeTGO 0.201914 0.117503 1.718 0.08617 .
## country_codeTZA -2.216959 0.288023 -7.697 4.75e-14 ***
## country_codeUGA -0.095200 0.162157 -0.587 0.55734
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3709 on 700 degrees of freedom
## Multiple R-Squared: 0.3391, Adjusted R-squared: 0.2995
## Wald test: 13.55 on 42 and 700 DF, p-value: < 2.2e-16
```

The coefficient on `LaggedGDPGrowth` has increased by 40% (-2 to -2.8) and now has a p-value of 0.06. These regressions use only variation within each country and thus remove the country-level unobserved variables that may be correlated with economic growth.

f. Regression `LaggedGDPGrowth` on rainfall and controls:

```
mod_f1 <- lm(LaggedGDPGrowth ~ LaggedRainfallGrowth + InitialGDP + Democracy + Mountains + EthnicFrac +
```

Save the residuals. I use `broom::augment` for convenience, but still need to add `InternalConflict` since it wasn't in the original data.

```
data_f <- broom::augment(mod_f1) %>%
  left_join(select(RainIV, country_code, InternalConflict), by = "country_code")

mod_f2 <- lm(InternalConflict ~ .resid + LaggedGDPGrowth + InitialGDP + Democracy + Mountains + EthnicFrac +
  coef(mod_f2)["LaggedGDPGrowth"]
```

```
## LaggedGDPGrowth
## -6.536277e-14
```

The coefficient is the same as that in the 2SLS.

It controls for endogeneity by controlling for the part of `LaggedGDPGrowth` not explained by rainfall growth (i.e. the non-random part of GDP growth).

## 0.2 Problem 2

Bailey (2016) Ex 9.2

## 0.3 Problem 3

Bailey (2016) Ex 9.4

## 0.4 Problem 4

Bailey (2016) Ex 9.5

## References

Bailey, Michael A. 2016. *Real Stats: Using Econometrics for Political Science and Public Policy*. Oxford University Press.