POLS/CS&SS 503: Assignment 3

Jeffrey Arnold

May 17, 2017

Setup

```
library("tidyverse")
library("broom")
library("haven")
library("AER")
library("sandwich")
library("Formula")
```

1 Problem 1

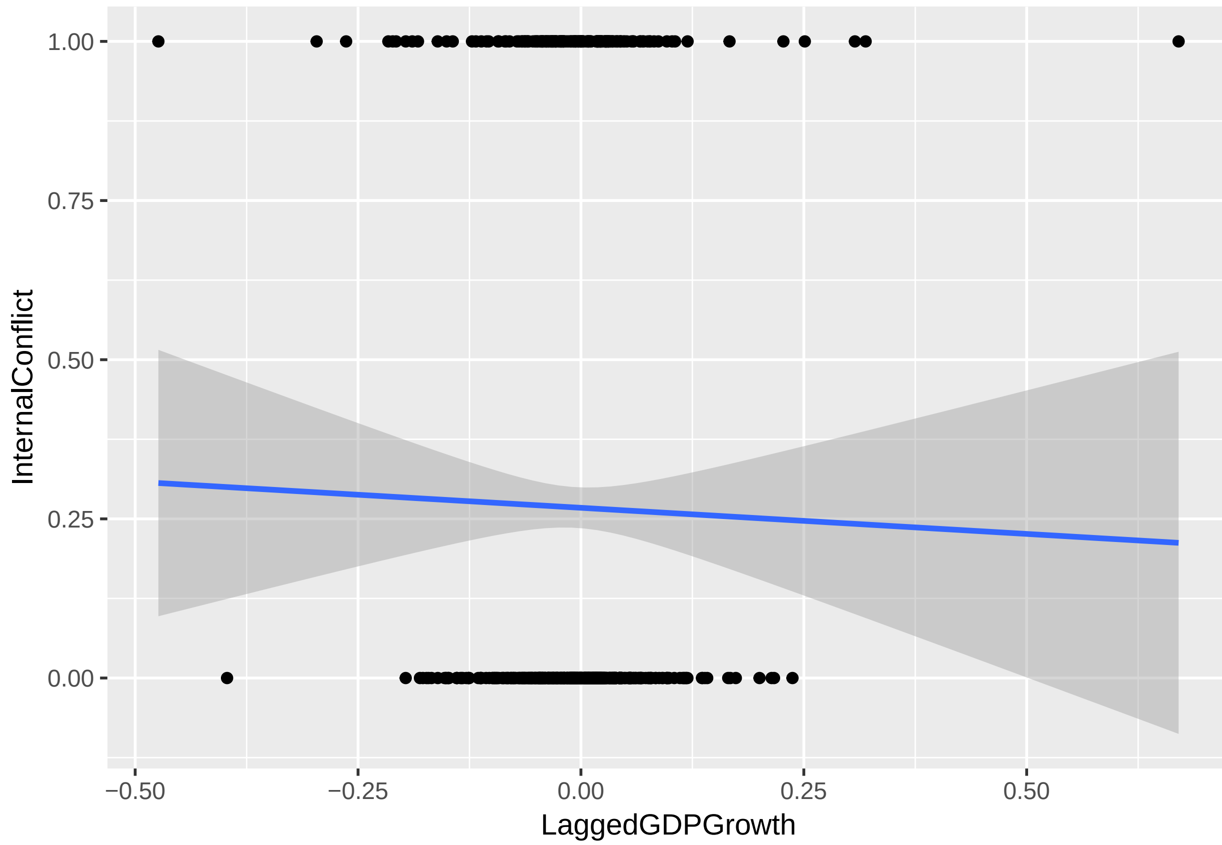
Bailey (2016) Ex 9.1

```
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")
```




```r
f_1a <- InternalConflict ~ LaggedGDPGrowth
mod_1a <- lm(f_1a, data = RainIV)
summary(mod_1a)
```

```
##
## Call:
## lm(formula = f_1a, 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.

```r
f_1b <- update(f_1a, . ~ . + InitialGDP + Democracy + Mountains + EthnicFrac +
                 ReligiousFrac)
```

```r
mod_1b <- lm(f_1b, data = RainIV)
summary(mod_1b)
```

```
##
## Call:
## lm(formula = f_1b, 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
```

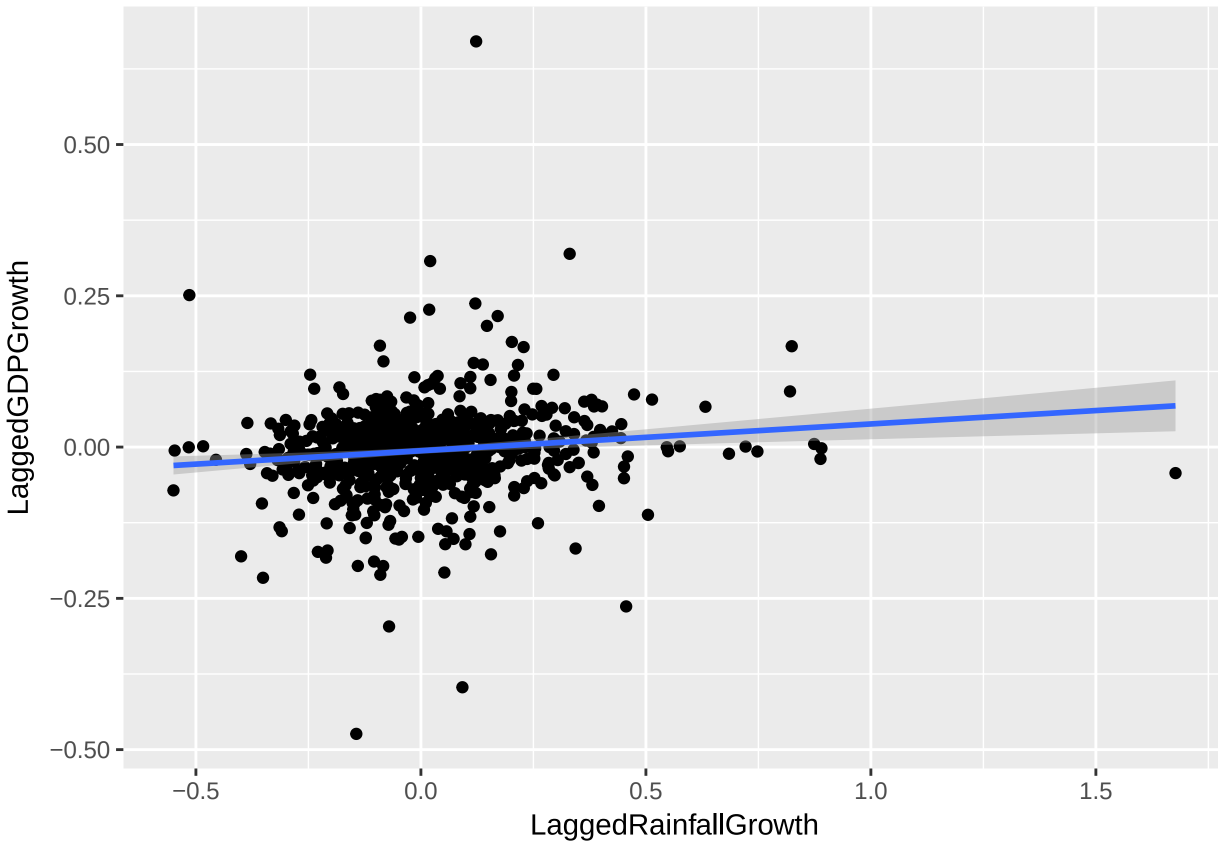
No. The coefficient on β would only be a causal estimate under an assumption of no omitted confounders. This seems highly unlikely.

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 correled with war excepct through its effect on economic growth.

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

```r
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.

```r
f_1e <- InternalConflict ~ LaggedGDPGrowth + InitialGDP + Democracy + Mountains + EthnicFrac + Religious
mod_1e <- ivreg(f_1e, data = RainIV)
summary(mod_1e)
```

```
##
## Call:
## ivreg(formula = f_1e, 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 stastistically significant.

f. Redo the 2SLS with country fixed effects.

```r
f_1f <- InternalConflict ~ LaggedGDPGrowth + InitialGDP + Democracy + Mountains + EthnicFrac + Religiou
mod_1f <- ivreg(f_1f, data = RainIV)
summary(mod_1f)
```

```
##
## Call:
## ivreg(formula = f_1f, 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.

g. Regression `LaggedGDPGrowth` on rainfall and controls:

```
mod_1g <- 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.

```
RainIV$resid <- residuals(mod_1g)
```

```
mod_1g <- lm(InternalConflict ~ LaggedGDPGrowth + resid + InitialGDP +
               Democracy + Mountains + EthnicFrac + ReligiousFrac + country_code,
             data = RainIV)
coef(mod_1g)["LaggedGDPGrowth"]
```

```
## LaggedGDPGrowth
##        -2.85338
```

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

```
coef(mod_1f)["LaggedGDPGrowth"]
```

```
## LaggedGDPGrowth
##        -2.85338
```

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).

2 Problem 2

Bailey (2016) Ex 9.2

```
NEWS_STUDY_FILE <- "Ch09.Ex2.TelevisionExperimentData/news_study_MAB.csv"
col_types <- cols(
  resid = col_character(),
  Female = col_character(),
```

```r
  watchnat = col_character(),
  ReadNews = col_character(),
  pnintst = col_character(),
  Education = col_character(),
  income = col_character(),
  Voted = col_character(),
  prop = col_character(),
  infopro = col_character(),
  WatchProgram = col_character(),
  learnpro = col_character(),
  TreatmentGroup = col_character(),
  prop_dv = col_integer(),
  InformationLevel = col_integer(),
  prop_vote = col_integer(),
  white = col_integer(),
  partyid = col_integer(),
  PoliticalInterest = col_character()
)
news_study <- read_csv(NEWS_STUDY_FILE, col_types = col_types) %>%
  mutate(WatchProgram = if_else(WatchProgram == "yes", 1, 0),
         TreatmentGroup = if_else(WatchProgram == "0", 0, 1))
```

Estimate a regression with Proposition 209 as a dependent variable and whether the person watched the program as the independent variable,

```r
mod_news_study_a <- lm(InformationLevel ~ WatchProgram, data = news_study)
coeftest(mod_news_study_a, vcov. = vcovHC(mod_news_study_a))["WatchProgram", ]
```

```
##     Estimate   Std. Error      t value     Pr(>|t|)
## 0.2963682432 0.0763601642 3.8811891825 0.0001179926
```

Those who watched the television program on average report 0.3 higher points ($p < .001$) on the information about Proposition 209. This should not be interpreted causally, since it is almost certainly biased due to endogeneity. One plausible example is that those more informed about politics are more likely to watch this TV program.

b. This regression controls for political interest, newspaper reading, and education.

```r
mod_news_study_b <-
  update(mod_news_study_a,
         . ~ . + PoliticalInterest + Education + ReadNews)
coeftest(mod_news_study_b, vcov. = vcovHC(mod_news_study_b))["WatchProgram", ]
```

```
##   Estimate Std. Error    t value   Pr(>|t|)
## 0.19261624 0.07613540 2.52991690 0.01173457
```

The result is smaller 0.19 vs. 0.3 with a higher p-value ($p < 0.05$). The identification strategy is selection on observables, which requires that all relevant variables are controlled for. This is unlikely.

c. The assignment variable is a good instrument since it is an experiment, so by construction it is not correlated with the dependent variable and thus not associated with `PoliticalInterest` except watching the TV show.

```r
mod_news_study_c <- lm(WatchProgram ~ TreatmentGroup + PoliticalInterest +
                         Education + ReadNews, data = news_study)

coeftest(mod_news_study_c,
         vcov. = vcovHC(mod_news_study_c))["TreatmentGroup", ]
```

```
##     Estimate   Std. Error      t value     Pr(>|t|)
## 1.000000e+00 4.088662e-16 2.445788e+15 0.000000e+00
```

d. Estimate a 2SLS using `TreatmentGroup` as an instrument for `WatchProgram`:

```
mod_news_study_d <-
  ivreg(InformationLevel ~ WatchProgram + PoliticalInterest +
          Education + ReadNews |
           . - WatchProgram + TreatmentGroup,
        data = news_study)

coeftest(mod_news_study_d)["WatchProgram", ]
```

```
##   Estimate Std. Error    t value   Pr(>|t|)
## 0.19261624 0.07700970 2.50119458 0.01271667
```

```
coeftest(mod_news_study_d, vcov. = vcovHC(mod_news_study_d))["WatchProgram", ]
```

```
##   Estimate Std. Error    t value   Pr(>|t|)
## 0.19261624 0.07613540 2.52991690 0.01173457
```

e. The 2SLS suggest that we can't reject that there is an effect of watching the program on information levels. Conditional on the IV assumptions (especially the inclusion restriction), this has removed endogeneity.

3 Problem 3

Bailey (2016) Ex 9.4

```
zipfile <- "Ch09.Ex2.TelevisionExperimentData.zip"
URL <- paste0("http://global.oup.com/",
              "us/companion.websites/fdscontent/uscompanion/us/",
              "static/companion.websites/9780199981946/data_sets/ch9/",
"Ch09.Ex4.EducationCrimeData.zip")
download.file(URL, destfile = zipfile)
unzip(zipfile)
```

```
EducationCrimeFile <- "Ch09.Ex4.EducationCrimeData/inmates.csv"
col_types <- cols(
  age = col_integer(),
  state = col_integer(),
  pob = col_integer(),
  gqtype = col_integer(),
  prison = col_integer(),
  educ = col_integer(),
  drop = col_integer(),
  AfAm = col_integer(),
  yearat14 = col_integer(),
  birthpl = col_integer(),
  req_sch = col_integer(),
  work_age = col_integer(),
  work_sch = col_integer(),
  ca = col_integer(),
  enrolage = col_integer(),
```

```r
  drop_age = col_integer(),
  cl = col_integer(),
  ca8 = col_integer(),
  ca9 = col_integer(),
  ca10 = col_integer(),
  ca11 = col_integer(),
  cl6 = col_integer(),
  cl7 = col_integer(),
  cl8 = col_integer(),
  cl9 = col_integer(),
  year = col_integer(),
  hisp = col_character()
)
EducationCrime <- read_csv(EducationCrimeFile, col_types = col_types)
```

a. Run a LPM with prison as the dependent variable

```r
formula_ex3a <- prison ~ educ + age + AfAm + factor(year) + factor(state)
mod_ex3a <- lm(formula_ex3a, data = EducationCrime)
#coeftest(mod_ex3a, vcov. = vcovHC)["educ", ]
tidy(mod_ex3a) %>% filter(term == "educ")
```

```
##   term     estimate    std.error statistic p.value
## 1 educ -0.001198227 1.391285e-05 -86.12376       0
```

On average those with one more year of schooling are 0.1% percent less likely to commit a crime, holding age and race constant ($p < 0.001$).

b. No. This depends on selection on observables. It is likely there is an unobservable that means that those who are more likely to complete schooling are less likely to commit crimes (SES, aggressiveness, conscientiousness).

c. Compulsary education laws should clearly be associated with more years of schooling, satisfying the inclusion restriction.

```r
mod_ex3c <- lm(educ ~ ca9 + ca10 + ca11 + age + AfAm + factor(state) +
                    factor(year), data = EducationCrime)
mod_ex3c_un <- lm(educ ~ age + AfAm + factor(state) + factor(year),
                     data = EducationCrime)
anova(mod_ex3c, mod_ex3c_un)
```

```
## Analysis of Variance Table
##
## Model 1: educ ~ ca9 + ca10 + ca11 + age + AfAm + factor(state) + factor(year)
## Model 2: educ ~ age + AfAm + factor(state) + factor(year)
##    Res.Df      RSS Df Sum of Sq      F    Pr(>F)
## 1 3610609 33600250
## 2 3610612 33743214 -3   -142964 5120.9 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
tidy(mod_ex3c) %>% filter(term %in% c("ca9", "ca10", "ca11"))
```

```
##   term  estimate   std.error statistic       p.value
## 1  ca9 0.4981164 0.005026353  99.10097  0.000000e+00
## 2 ca10 0.2735032 0.007207631  37.94634 5.119031e-315
## 3 ca11 0.6498893 0.005627761 115.47918  0.000000e+00
```

It's not as clear that it satisfies the exclusion restriction. It is plausible that states with more crime are more likely to pass compulsary schooling laws.

d. A 2SLS model using these instruments and robust se:

```
mod_ex3d <- ivreg(prison ~ educ + age + AfAm + factor(state) + factor(year) |
                  . - educ + ca9 + ca10 + ca11, data = EducationCrime)
# coeftest(mod_ex3d, vcov. = vcovHC)
summary(mod_ex3d)
```

```
##
## Call:
## ivreg(formula = prison ~ educ + age + AfAm + factor(state) +
##     factor(year) | . - educ + ca9 + ca10 + ca11, data = EducationCrime)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -0.0510583 -0.0097271 -0.0050842 -0.0005812  1.0144415
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)      2.983e-02  2.577e-03  11.575  < 2e-16 ***
## educ            -1.259e-03  2.137e-04  -5.888 3.91e-09 ***
## age             -3.778e-04  1.127e-05 -33.525  < 2e-16 ***
## AfAm             2.108e-02  3.888e-04  54.202  < 2e-16 ***
## factor(state)2  -1.664e-05  1.105e-03  -0.015 0.987982
## factor(state)4   5.219e-03  5.496e-04   9.495  < 2e-16 ***
## factor(state)5   2.438e-04  5.376e-04   0.453 0.650213
## factor(state)6   5.550e-03  4.584e-04  12.109  < 2e-16 ***
## factor(state)8   3.109e-03  5.639e-04   5.513 3.52e-08 ***
## factor(state)9   2.260e-03  5.489e-04   4.118 3.83e-05 ***
## factor(state)10  2.824e-03  8.924e-04   3.165 0.001553 **
## factor(state)11 -8.074e-03  9.047e-04  -8.925  < 2e-16 ***
## factor(state)12  6.031e-03  4.183e-04  14.418  < 2e-16 ***
## factor(state)13  3.756e-03  4.178e-04   8.991  < 2e-16 ***
## factor(state)15 -2.250e-03  1.117e-03  -2.013 0.044079 *
## factor(state)16  3.306e-03  7.678e-04   4.306 1.66e-05 ***
## factor(state)17  1.191e-03  4.200e-04   2.836 0.004566 **
## factor(state)18  2.747e-03  4.265e-04   6.440 1.19e-10 ***
## factor(state)19  1.991e-03  5.082e-04   3.918 8.93e-05 ***
## factor(state)20  6.286e-03  5.571e-04  11.281  < 2e-16 ***
## factor(state)21  2.221e-03  4.763e-04   4.662 3.13e-06 ***
## factor(state)22  2.322e-03  4.461e-04   5.205 1.94e-07 ***
## factor(state)23  1.551e-03  6.859e-04   2.261 0.023775 *
## factor(state)24  2.420e-03  4.834e-04   5.007 5.53e-07 ***
## factor(state)25  1.892e-03  4.872e-04   3.883 0.000103 ***
## factor(state)26  2.543e-03  4.091e-04   6.215 5.15e-10 ***
## factor(state)27  2.223e-03  4.859e-04   4.575 4.77e-06 ***
## factor(state)28 -3.270e-03  5.190e-04  -6.301 2.96e-10 ***
## factor(state)29  2.405e-03  4.423e-04   5.438 5.40e-08 ***
## factor(state)30  2.852e-03  7.931e-04   3.595 0.000324 ***
## factor(state)31  2.441e-03  6.254e-04   3.904 9.47e-05 ***
## factor(state)32  5.021e-03  8.201e-04   6.122 9.22e-10 ***
## factor(state)33  6.735e-04  7.629e-04   0.883 0.377369
## factor(state)34  1.394e-03  4.671e-04   2.984 0.002842 **
```

```
## factor(state)35  2.642e-03  6.668e-04   3.962 7.45e-05 ***
## factor(state)36  2.787e-03  4.378e-04   6.366 1.94e-10 ***
## factor(state)37  2.434e-03  4.097e-04   5.941 2.83e-09 ***
## factor(state)38  2.786e-04  8.354e-04   0.333 0.738785
## factor(state)39  2.492e-03  3.934e-04   6.335 2.38e-10 ***
## factor(state)40  4.521e-03  5.052e-04   8.949  < 2e-16 ***
## factor(state)41  4.340e-03  5.563e-04   7.802 6.10e-15 ***
## factor(state)42  1.532e-03  3.902e-04   3.928 8.57e-05 ***
## factor(state)44  3.702e-04  7.388e-04   0.501 0.616291
## factor(state)45  1.857e-04  4.793e-04   0.387 0.698388
## factor(state)46  2.544e-03  8.282e-04   3.071 0.002131 **
## factor(state)47  1.653e-03  4.355e-04   3.795 0.000148 ***
## factor(state)48  4.781e-03  3.810e-04  12.550  < 2e-16 ***
## factor(state)49  2.808e-03  7.071e-04   3.972 7.13e-05 ***
## factor(state)50  1.554e-03  9.643e-04   1.611 0.107121
## factor(state)51  3.353e-03  4.231e-04   7.925 2.27e-15 ***
## factor(state)53  4.348e-03  5.124e-04   8.486  < 2e-16 ***
## factor(state)54  5.394e-04  5.573e-04   0.968 0.333103
## factor(state)55  1.980e-03  4.553e-04   4.349 1.36e-05 ***
## factor(state)56  1.968e-03  9.796e-04   2.009 0.044528 *
## factor(year)70  -2.666e-04  2.693e-04  -0.990 0.322059
## factor(year)80   1.679e-03  4.419e-04   3.800 0.000145 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.08082 on 3610611 degrees of freedom
## Multiple R-Squared: 0.0128,  Adjusted R-squared: 0.01278
## Wald test: 717.1 on 55 and 3610611 DF,  p-value: < 2.2e-16
```

4 Problem 4

Bailey (2016) Ex 9.5

```
GrowthDemocracyFile <- "Ch09.Ex5.GrowthDemocracyData/democracy_income.csv"
col_types = cols(
  CountryCode = col_integer(),
  democracy_fh = col_double(),
  log_gdp = col_double(),
  year = col_integer(),
  worldincome = col_double(),
  YearOrder = col_integer()
)
GrowthDemocracy <- read_csv(GrowthDemocracyFile, col_types = col_types) %>%
  # add lag GDP
  group_by(CountryCode) %>%
  arrange(CountryCode, year) %>%
  mutate(lag_log_gdp = lag(log_gdp),
         lag_worldincome = lag(worldincome))
```

a. Run a model with Democracy as the dependent variable and logged GDP per capita as a

```
mod_ex5a <- lm(democracy_fh ~ lag_log_gdp, data = GrowthDemocracy)
tidy(mod_ex5a) %>% filter(term == "lag_log_gdp")
```

```
##          term  estimate   std.error statistic      p.value
## 1 lag_log_gdp 0.2337698 0.008984252  26.01995 7.47271e-112
```

b. Include fixed effects of year and country in the previous model,

```
mod_ex5b <- lm(democracy_fh ~ lag_log_gdp + factor(year) + factor(CountryCode),
   data = GrowthDemocracy)
tidy(mod_ex5b) %>% filter(term == "lag_log_gdp")
```

```
##          term   estimate  std.error statistic   p.value
## 1 lag_log_gdp 0.03840844 0.02899967  1.324444 0.1857472
```

This uses only variation within country and within year. This explains much of the variation in democracy—e.g. time trends and variables constant within country.

c. World income of trading partners should be associated with GDP. It is both plausible, and has a t-stat of > 3 in the 1st stage regression.

```
mod_ex5c <- lm(log_gdp ~ worldincome + factor(year) + factor(CountryCode),
               data = GrowthDemocracy)
tidy(mod_ex5c) %>% filter(term == "worldincome")
```

```
##          term  estimate  std.error statistic      p.value
## 1 worldincome 0.4074482 0.05081291  8.018596 3.395804e-15
```

The exclusion restriction is not entirely plausible, but I can't think of a strong reason to object. Controlling for year is important since it will control global trade shocks. What would be problematic is local changes which increase both democracy and GDP per capita – since most trading partners are close neighbors it is plausible to think of spillovers. However, shocks to GDP in the short run seem more plausible than shocks to democracy.

d. Run 2SLS with `worldincome` as an instrument for logged GDP,

```
mod_ex5d <- ivreg(democracy_fh ~ lag_log_gdp +
                    factor(year) + factor(CountryCode) |
                  . - lag_log_gdp + lag_worldincome,
                  data = GrowthDemocracy)
summary(mod_ex5d)
```

```
##
## Call:
## ivreg(formula = democracy_fh ~ lag_log_gdp + factor(year) + factor(CountryCode) |
##     . - lag_log_gdp + lag_worldincome, data = GrowthDemocracy)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.740829 -0.098046 -0.003134  0.103748  0.588676
##
## Coefficients:
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)             1.511634   0.795809   1.899 0.057872 .
## lag_log_gdp            -0.212988   0.116458  -1.829 0.067802 .
## factor(year)1965        0.052065   0.042990   1.211 0.226231
## factor(year)1970       -0.032971   0.054500  -0.605 0.545371
## factor(year)1975       -0.019955   0.066164  -0.302 0.763044
## factor(year)1980        0.058245   0.077258   0.754 0.451134
## factor(year)1985        0.090943   0.085582   1.063 0.288273
## factor(year)1990        0.141965   0.088543   1.603 0.109269
## factor(year)1995        0.214055   0.095932   2.231 0.025945 *
```

```
## factor(year)2000        0.244586   0.100265   2.439 0.014936 *
## factor(CountryCode)6    1.000968   0.236304   4.236 2.55e-05 ***
## factor(CountryCode)8    0.843134   0.262336   3.214 0.001364 **
## factor(CountryCode)9    1.455676   0.295719   4.922 1.04e-06 ***
## factor(CountryCode)10   1.402012   0.272214   5.150 3.30e-07 ***
## factor(CountryCode)14  -0.155311   0.151585  -1.025 0.305883
## factor(CountryCode)16   0.138747   0.125361   1.107 0.268736
## factor(CountryCode)17   0.064932   0.144226   0.450 0.652685
## factor(CountryCode)18   0.373538   0.139387   2.680 0.007522 **
## factor(CountryCode)24   1.117314   0.193426   5.776 1.11e-08 ***
## factor(CountryCode)25   0.633451   0.133523   4.744 2.49e-06 ***
## factor(CountryCode)26   0.832586   0.168191   4.950 9.10e-07 ***
## factor(CountryCode)27   1.327520   0.239295   5.548 3.98e-08 ***
## factor(CountryCode)30   0.950612   0.144823   6.564 9.60e-11 ***
## factor(CountryCode)31   0.213614   0.117809   1.813 0.070185 .
## factor(CountryCode)32   1.464057   0.299949   4.881 1.28e-06 ***
## factor(CountryCode)33   1.512442   0.324554   4.660 3.72e-06 ***
## factor(CountryCode)34   0.858879   0.181323   4.737 2.58e-06 ***
## factor(CountryCode)35  -0.027954   0.120092  -0.233 0.816001
## factor(CountryCode)36   0.226918   0.123151   1.843 0.065771 .
## factor(CountryCode)37   0.183033   0.120100   1.524 0.127915
## factor(CountryCode)38   0.114527   0.120128   0.953 0.340696
## factor(CountryCode)39   0.861877   0.155637   5.538 4.20e-08 ***
## factor(CountryCode)40   0.300133   0.131037   2.290 0.022264 *
## factor(CountryCode)41   0.512517   0.132153   3.878 0.000114 ***
## factor(CountryCode)42   1.168539   0.171188   6.826 1.76e-11 ***
## factor(CountryCode)43   0.146808   0.203992   0.720 0.471942
## factor(CountryCode)44   1.131723   0.207296   5.459 6.44e-08 ***
## factor(CountryCode)47   1.388909   0.309418   4.489 8.25e-06 ***
## factor(CountryCode)51   1.045899   0.180208   5.804 9.45e-09 ***
## factor(CountryCode)52   1.466359   0.301112   4.870 1.36e-06 ***
## factor(CountryCode)53   0.698774   0.127221   5.493 5.38e-08 ***
## factor(CountryCode)54   0.358677   0.163373   2.195 0.028428 *
## factor(CountryCode)55   0.692357   0.140147   4.940 9.57e-07 ***
## factor(CountryCode)56   0.325371   0.122716   2.651 0.008180 **
## factor(CountryCode)58   0.980129   0.241409   4.060 5.41e-05 ***
## factor(CountryCode)60  -0.087582   0.272763  -0.321 0.748228
## factor(CountryCode)61  -0.091143   0.156731  -0.582 0.561055
## factor(CountryCode)63   1.333144   0.274798   4.851 1.48e-06 ***
## factor(CountryCode)64   0.707605   0.165011   4.288 2.03e-05 ***
## factor(CountryCode)65   1.399007   0.277899   5.034 5.97e-07 ***
## factor(CountryCode)66   0.559385   0.207746   2.693 0.007242 **
## factor(CountryCode)67   1.423045   0.280466   5.074 4.89e-07 ***
## factor(CountryCode)70   0.229677   0.123718   1.856 0.063769 .
## factor(CountryCode)71   0.171455   0.129842   1.320 0.187063
## factor(CountryCode)72   0.469163   0.125878   3.727 0.000208 ***
## factor(CountryCode)73   0.032907   0.182134   0.181 0.856670
## factor(CountryCode)74   0.025857   0.126650   0.204 0.838282
## factor(CountryCode)75   1.141115   0.229850   4.965 8.47e-07 ***
## factor(CountryCode)76   0.999285   0.164919   6.059 2.14e-09 ***
## factor(CountryCode)77   0.682909   0.143910   4.745 2.48e-06 ***
## factor(CountryCode)78   0.585027   0.131038   4.465 9.22e-06 ***
## factor(CountryCode)79   0.559226   0.119443   4.682 3.36e-06 ***
## factor(CountryCode)81   0.065748   0.138077   0.476 0.634085
```

```
## factor(CountryCode)82   0.859921   0.215982   3.981 7.50e-05 ***
## factor(CountryCode)83   0.257445   0.117357   2.194 0.028554 *
## factor(CountryCode)84   0.724406   0.117773   6.151 1.24e-09 ***
## factor(CountryCode)85   1.324210   0.235375   5.626 2.58e-08 ***
## factor(CountryCode)86   0.431779   0.158647   2.722 0.006642 **
## factor(CountryCode)88   1.432493   0.284077   5.043 5.73e-07 ***
## factor(CountryCode)89   1.250901   0.244638   5.113 4.00e-07 ***
## factor(CountryCode)90   1.395002   0.268742   5.191 2.68e-07 ***
## factor(CountryCode)91   1.017408   0.152158   6.687 4.38e-11 ***
## factor(CountryCode)92   0.405446   0.137113   2.957 0.003201 **
## factor(CountryCode)93   1.346546   0.260377   5.172 2.96e-07 ***
## factor(CountryCode)95   0.093697   0.129199   0.725 0.468541
## factor(CountryCode)99   1.206026   0.227188   5.309 1.45e-07 ***
## factor(CountryCode)101  0.721934   0.159830   4.517 7.26e-06 ***
## factor(CountryCode)107  1.105865   0.172242   6.420 2.37e-10 ***
## factor(CountryCode)109  0.738260   0.117143   6.302 4.93e-10 ***
## factor(CountryCode)110  0.238834   0.131017   1.823 0.068701 .
## factor(CountryCode)114  0.458021   0.130115   3.520 0.000457 ***
## factor(CountryCode)116  0.399247   0.125264   3.187 0.001494 **
## factor(CountryCode)118  0.848495   0.197093   4.305 1.88e-05 ***
## factor(CountryCode)120  0.104768   0.137218   0.764 0.445390
## factor(CountryCode)125  0.148302   0.144998   1.023 0.306732
## factor(CountryCode)126  0.111595   0.117148   0.953 0.341094
## factor(CountryCode)127  1.125612   0.198551   5.669 2.03e-08 ***
## factor(CountryCode)128  0.007338   0.170259   0.043 0.965632
## factor(CountryCode)129  0.767469   0.162987   4.709 2.95e-06 ***
## factor(CountryCode)130  0.872113   0.194874   4.475 8.78e-06 ***
## factor(CountryCode)131  0.110271   0.121237   0.910 0.363342
## factor(CountryCode)132  0.190609   0.125162   1.523 0.128191
## factor(CountryCode)133  0.596319   0.142000   4.199 2.99e-05 ***
## factor(CountryCode)134  1.435485   0.285574   5.027 6.21e-07 ***
## factor(CountryCode)135  1.431166   0.283413   5.050 5.52e-07 ***
## factor(CountryCode)136  0.321585   0.132232   2.432 0.015243 *
## factor(CountryCode)137  1.438848   0.287260   5.009 6.79e-07 ***
## factor(CountryCode)140  0.250041   0.132598   1.886 0.059709 .
## factor(CountryCode)141  0.364770   0.177988   2.049 0.040759 *
## factor(CountryCode)142  0.668302   0.161847   4.129 4.04e-05 ***
## factor(CountryCode)144  0.755602   0.166442   4.540 6.53e-06 ***
## factor(CountryCode)145  0.725007   0.131216   5.525 4.50e-08 ***
## factor(CountryCode)147  0.892721   0.146923   6.076 1.93e-09 ***
## factor(CountryCode)148  0.946058   0.205683   4.600 4.94e-06 ***
## factor(CountryCode)150  0.982876   0.211303   4.651 3.88e-06 ***
## factor(CountryCode)151  0.568335   0.152717   3.721 0.000212 ***
## factor(CountryCode)153  0.290625   0.133371   2.179 0.029628 *
## factor(CountryCode)155 -0.088428   0.138508  -0.638 0.523381
## factor(CountryCode)160  0.404959   0.120453   3.362 0.000812 ***
## factor(CountryCode)162  0.724986   0.239171   3.031 0.002517 **
## factor(CountryCode)165  0.191632   0.124626   1.538 0.124543
## factor(CountryCode)166  0.838298   0.159998   5.239 2.08e-07 ***
## factor(CountryCode)168  0.339155   0.139598   2.430 0.015346 *
## factor(CountryCode)172  1.434579   0.294420   4.873 1.34e-06 ***
## factor(CountryCode)174  0.652452   0.222295   2.935 0.003434 **
## factor(CountryCode)175  0.165854   0.128643   1.289 0.197695
## factor(CountryCode)176  0.034181   0.123085   0.278 0.781318
```

```
## factor(CountryCode)177  0.056079   0.122300   0.459 0.646696
## factor(CountryCode)178  0.602021   0.124609   4.831 1.64e-06 ***
## factor(CountryCode)182  1.207397   0.219898   5.491 5.43e-08 ***
## factor(CountryCode)183  0.358918   0.154938   2.317 0.020790 *
## factor(CountryCode)184  0.780971   0.162593   4.803 1.88e-06 ***
## factor(CountryCode)187 -0.031677   0.176244  -0.180 0.857407
## factor(CountryCode)188 -0.051740   0.165546  -0.313 0.754713
## factor(CountryCode)191  1.058565   0.210854   5.020 6.41e-07 ***
## factor(CountryCode)192  1.482195   0.316485   4.683 3.33e-06 ***
## factor(CountryCode)195  0.987634   0.174491   5.660 2.13e-08 ***
## factor(CountryCode)196  1.146777   0.228673   5.015 6.58e-07 ***
## factor(CountryCode)197 -0.206241   0.177090  -1.165 0.244537
## factor(CountryCode)208  0.898084   0.209955   4.278 2.13e-05 ***
## factor(CountryCode)209 -0.111803   0.146543  -0.763 0.445734
## factor(CountryCode)210  0.250349   0.126633   1.977 0.048401 *
## factor(CountryCode)211  0.354456   0.127761   2.774 0.005665 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2052 on 773 degrees of freedom
## Multiple R-Squared: 0.726,   Adjusted R-squared: 0.6792
## Wald test: 16.05 on 132 and 773 DF,  p-value: < 2.2e-16
```

The coefficient switches signs from 0.38 in the panel data to -0.21 in `ivreg` and the p-value decreases from 0.18 to 0.06.

It would probably be better to run this with cluster robust standard errors.

References

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