Assignment three

Christianna Parr

May 19, 2017

Chapter 9: Instrumental Variables. Assignment Three: Q 1, 2, 4, and 5.

```r
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
```

```r
library (broom)
library(rio)
library (modelr)
```

```
##
## Attaching package: 'modelr'

## The following object is masked from 'package:broom':
##
##     bootstrap
```

```r
library (AER)
```

```
## Warning: package 'AER' was built under R version 3.3.3

## Loading required package: car

## Warning: package 'car' was built under R version 3.3.3

##
## 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

## Warning: package 'lmtest' was built under R version 3.3.3

## Loading required package: zoo

## Warning: package 'zoo' was built under R version 3.3.3
```

```
##
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric

## Loading required package: sandwich

## Warning: package 'sandwich' was built under R version 3.3.3

## Loading required package: survival
```

```
library (car)
library (plm)
```

```
## Warning: package 'plm' was built under R version 3.3.3

## Loading required package: Formula

##
## Attaching package: 'plm'

## The following objects are masked from 'package:dplyr':
##
##     between, lag, lead
```

```
rain <- read_csv ("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.
```

1 a)

```
# 1. a. Estimate a bivariate OLS model with internal conflict and GDP.

ols <- tidy(lm (InternalConflict ~ LaggedGDPGrowth, data = rain))
ols
```

```
##              term    estimate  std.error  statistic      p.value
## 1     (Intercept)  0.26737746 0.01631415 16.3892984 9.586655e-52
```

```
## 2 LaggedGDPGrowth -0.08206485 0.22485213 -0.3649725 7.152360e-01
```

When we run a bivariate OLS model we find no significant results (at the alpha equals 0.05) to suggest that growth in GDP would reduce conflict.

b)

```r
# b. with controls

olscontrol <- tidy(lm(InternalConflict ~ LaggedGDPGrowth + InitialGDP + Democracy + Mountains + EthnicF
olscontrol
```

```
##              term     estimate    std.error  statistic      p.value
## 1     (Intercept)  0.070355529 0.0731012386  0.9624396 3.361449e-01
## 2 LaggedGDPGrowth -0.108797693 0.2200998529 -0.4943106 6.212343e-01
## 3      InitialGDP -0.056909063 0.0182258230 -3.1224413 1.863801e-03
## 4       Democracy  0.001224162 0.0028894131  0.4236714 6.719292e-01
## 5       Mountains  0.003865434 0.0009526937  4.0573730 5.493161e-05
## 6      EthnicFrac  0.324793069 0.0918181328  3.5373521 4.295018e-04
## 7   ReligiousFrac  0.010516152 0.0958907296  0.1096681 9.127025e-01
```

These results also do not establish a causal relationship between economic growth and reduction in civil conflict (the p values are not significant at the 0.05 level).

c)

```r
#c. Instrumental variable test for rainfall

ols2 <- tidy(lm (LaggedGDPGrowth ~ LaggedRainfallGrowth + InitialGDP + Democracy + Mountains + EthnicFr

ols2
```

```
##                   term      estimate    std.error  statistic      p.value
## 1          (Intercept) -0.0058040646 0.0121558449 -0.4774711 0.6331684550
## 2 LaggedRainfallGrowth  0.0439769947 0.0128127339  3.4322881 0.0006319632
## 3           InitialGDP -0.0008055799 0.0030286889 -0.2659830 0.7903267540
## 4            Democracy  0.0005373436 0.0004797716  1.1199986 0.2630797024
## 5            Mountains  0.0001086811 0.0001582464  0.6867842 0.4924350511
## 6           EthnicFrac  0.0031602813 0.0152584101  0.2071173 0.8359755120
## 7        ReligiousFrac -0.0017176029 0.0159357476 -0.1077830 0.9141971916
```

The two conditions for a good instrument: 1. Inclusion condition: the instrument must explain X. 2. Exclusion condition: the instrument must explain x without explaining y.

We can only test for the first condition using a linear regression, like done in part c. Here we find the instrument explains x at an alpha of 0.05. But we cannot explain the exclusion condition. The only way to do this is through theory and justification. Rainfall can explain GDP growth but not conflict according to the authors of the study.

d.Why use rainfall as an instrument?

If we use rainfall as an instrument for GDP growth we can make a connection between the two, since rainfall would definitely effect agrarian societies. Larger amounts would signal more harvest and better export. However, rainfall would be a hard factor in thinking about conflict, other authors (Acemoglu and Robinson) do mention possible flooding of roads as a factor but dismiss this in their paper as unlikely. Therefore, rainfall can explain conflict but only through economic growth.

e)

```r
# e. Real Instrument test with ivreg

instrument <- ivreg(InternalConflict ~ LaggedGDPGrowth + InitialGDP + Democracy + Mountains + EthnicFra

summary (instrument)
```

```
##
## Call:
## ivreg(formula = InternalConflict ~ LaggedGDPGrowth + InitialGDP +
##     Democracy + Mountains + EthnicFrac + ReligiousFrac | LaggedRainfallGrowth +
##     InitialGDP + Democracy + Mountains + EthnicFrac + ReligiousFrac,
##     data = rain)
##
## 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
```

From this 2SLS we can see that there is not a significant result (at the alpha equals 0.05 level) for GDP growth even with the instrumental variable. We cannot reject the null hypothesis.

f)

```r
#f. Dummy variables

reg3 <- ivreg(InternalConflict ~ LaggedGDPGrowth + InitialGDP + Democracy + Mountains +
                EthnicFrac + ReligiousFrac + factor(country_name) | LaggedRainfallGrowth + InitialGDP +
                Democracy + Mountains + EthnicFrac + ReligiousFrac + factor(country_name), data = rain)

summary(reg3)
```

```
##
## Call:
```

```
## ivreg(formula = InternalConflict ~ LaggedGDPGrowth + InitialGDP +
##     Democracy + Mountains + EthnicFrac + ReligiousFrac + factor(country_name) |
##     LaggedRainfallGrowth + InitialGDP + Democracy + Mountains +
##         EthnicFrac + ReligiousFrac + factor(country_name), data = rain)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.60872 -0.18282 -0.01501  0.13649  1.92662
##
## Coefficients:
##                                                 Estimate Std. Error
## (Intercept)                                     0.459156   1.863122
## LaggedGDPGrowth                                -2.853380   1.535631
## InitialGDP                                     -0.476826   0.792632
## Democracy                                       0.001065   0.003276
## Mountains                                       0.092267   0.020715
## EthnicFrac                                     -0.177308   0.960665
## ReligiousFrac                                   0.122734   1.081919
## factor(country_name)Benin                       0.102932   0.151373
## factor(country_name)Botswana                    0.459522   0.379597
## factor(country_name)Burkina Faso                0.001897   0.456368
## factor(country_name)Burundi                    -6.759628   1.224003
## factor(country_name)Cameroon                   -1.449536   0.539017
## factor(country_name)Central African Republic   -0.527030   0.202258
## factor(country_name)Chad                       -0.049368   0.121434
## factor(country_name)Congo                       0.594180   0.520803
## factor(country_name)Djibouti                    0.108725   0.362860
## factor(country_name)Ethiopia                   -5.942561   1.157047
## factor(country_name)Gabon                       1.894540   2.969182
## factor(country_name)Gambia                      0.146595   0.339041
## factor(country_name)Ghana                       0.160302   0.216191
## factor(country_name)Guinea                     -0.165534   0.412092
## factor(country_name)Guinea-Bissau              -0.012907   0.290915
## factor(country_name)Ivory Coast                 0.570719   1.158194
## factor(country_name)Kenya                      -2.322296   0.572570
## factor(country_name)Lesotho                    -7.505690   1.488785
## factor(country_name)Liberia                     0.260105   0.270259
## factor(country_name)Madagascar                 -3.146483   0.563717
## factor(country_name)Malawi                     -0.962630   0.304634
## factor(country_name)Mali                       -0.011005   0.632748
## factor(country_name)Mauritania                  0.051883   0.865955
## factor(country_name)Mozambique                  0.502944   0.170638
## factor(country_name)Namibia                     0.191278   1.379937
## factor(country_name)Niger                       0.029188   0.386649
## factor(country_name)Nigeria                     0.016553   0.535843
## factor(country_name)Rwanda                     -6.447391   1.211757
## factor(country_name)Senegal                     0.581121   0.314757
## factor(country_name)Sierra Leone                0.384331   0.279397
## factor(country_name)Somalia                    -0.202700   0.986690
## factor(country_name)South Africa                1.195777   2.036813
## factor(country_name)Sudan                       0.352135   0.189502
## factor(country_name)Swaziland                  -0.336431   1.607692
## factor(country_name)Tanzania, United Republic of -2.107378   0.320915
## factor(country_name)Togo                        0.014390   0.326231
```

```
##                                                t value Pr(>|t|)
## (Intercept)                                      0.246  0.80541
## LaggedGDPGrowth                                 -1.858  0.06357 .
## InitialGDP                                      -0.602  0.54765
## Democracy                                        0.325  0.74518
## Mountains                                        4.454 9.81e-06 ***
## EthnicFrac                                      -0.185  0.85362
## ReligiousFrac                                    0.113  0.90971
## factor(country_name)Benin                        0.680  0.49674
## factor(country_name)Botswana                     1.211  0.22648
## factor(country_name)Burkina Faso                 0.004  0.99668
## factor(country_name)Burundi                     -5.523 4.71e-08 ***
## factor(country_name)Cameroon                    -2.689  0.00733 **
## factor(country_name)Central African Republic    -2.606  0.00936 **
## factor(country_name)Chad                        -0.407  0.68447
## factor(country_name)Congo                        1.141  0.25430
## factor(country_name)Djibouti                     0.300  0.76455
## factor(country_name)Ethiopia                    -5.136 3.64e-07 ***
## factor(country_name)Gabon                        0.638  0.52364
## factor(country_name)Gambia                       0.432  0.66560
## factor(country_name)Ghana                        0.741  0.45865
## factor(country_name)Guinea                      -0.402  0.68803
## factor(country_name)Guinea-Bissau               -0.044  0.96463
## factor(country_name)Ivory Coast                  0.493  0.62233
## factor(country_name)Kenya                       -4.056 5.56e-05 ***
## factor(country_name)Lesotho                     -5.041 5.89e-07 ***
## factor(country_name)Liberia                      0.962  0.33617
## factor(country_name)Madagascar                  -5.582 3.41e-08 ***
## factor(country_name)Malawi                      -3.160  0.00165 **
## factor(country_name)Mali                        -0.017  0.98613
## factor(country_name)Mauritania                   0.060  0.95224
## factor(country_name)Mozambique                   2.947  0.00331 **
## factor(country_name)Namibia                      0.139  0.88980
## factor(country_name)Niger                        0.075  0.93985
## factor(country_name)Nigeria                      0.031  0.97536
## factor(country_name)Rwanda                      -5.321 1.39e-07 ***
## factor(country_name)Senegal                      1.846  0.06528 .
## factor(country_name)Sierra Leone                 1.376  0.16939
## factor(country_name)Somalia                     -0.205  0.83729
## factor(country_name)South Africa                 0.587  0.55734
## factor(country_name)Sudan                        1.858  0.06356 .
## factor(country_name)Swaziland                   -0.209  0.83430
## factor(country_name)Tanzania, United Republic of  -6.567 1.00e-10 ***
## factor(country_name)Togo                         0.044  0.96483
## ---
## 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
```

In this 2SLS regression with dummy variables we find an effect at the 0.1 confidence level, but not at 0.05. Including the country fixed effects decreases the p value of the GDP growth (0.26 to 0.06). This shows us that when we control for confounders through dummies we can improve the certainty of our statistic.

G)

```r
# g) funky

firststage <- lm (LaggedGDPGrowth ~ LaggedRainfallGrowth + InitialGDP + Democracy + Mountains + EthnicFr
                        ReligiousFrac + country_code, data = rain)

rstage <- resid(firststage) #saving residuals

regl <- tidy(lm (InternalConflict ~ LaggedGDPGrowth + InitialGDP +Democracy + Mountains +EthnicFrac + Re
      country_name + rstage, data = rain))

head(regl)
```

```
##                term     estimate  std.error  statistic      p.value
## 1     (Intercept)  0.459155635 1.56002689  0.2943255 7.685966e-01
## 2 LaggedGDPGrowth -2.853379695 1.28581202 -2.2191266 2.679844e-02
## 3      InitialGDP -0.476825684 0.66368551 -0.7184513 4.727192e-01
## 4       Democracy  0.001065045 0.00274286  0.3882973 6.979142e-01
## 5       Mountains  0.092267006 0.01734527  5.3194329 1.402880e-07
## 6      EthnicFrac -0.177308059 0.80438281 -0.2204275 8.256026e-01
```

Correction: The coefficients are the same (-2.8) when we run this model. It is controlling for endogeneity since it controls for other things that rainfall cannot.

Question 2:

```r
tv <- import ("news_study_MAB.dta")
```

a) Bivariate OLS

```r
reg1 <- tidy(lm (InformationLevel ~ WatchProgram, data = tv))

reg1
```

```
##           term  estimate  std.error statistic       p.value
## 1  (Intercept) 3.1567568 0.04270106 73.926887 5.884937e-270
## 2 WatchProgram 0.2963682 0.08422643  3.518708  4.735592e-04
```

The results can be biased because there may be a selection problem, those who were likely to watch the tv program might have had previous higher levels of information. There could be a problem with causality in this situation.

b) With controls

```r
reg2 <- tidy(lm (InformationLevel ~ WatchProgram + PoliticalInterest + ReadNews + Education, data = tv)

reg2
```

```
##                term     estimate  std.error   statistic      p.value
## 1       (Intercept) 1.6942655996 0.16400904 10.33031811 9.323271e-23
```

```
## 2      WatchProgram 0.2329058838 0.07695731  3.02642967 2.605338e-03
## 3 PoliticalInterest 0.2650756301 0.04600878  5.76141449 1.478782e-08
## 4          ReadNews 0.1087892561 0.01827182  5.95393810 5.008485e-09
## 5         Education 0.0008843813 0.01242477  0.07117888 9.432846e-01
```

The results do not appear to be as different from each other. The coefficients are relatively the same, only reducing slightly in the controlled regression. The estimate for watching the program is still significant. But there is a problem with endogeneity because multiple variables could be corelated with the error term. The causal direction is still uncertain.

c) Instrument test

```r
reg3 <- lm (WatchProgram ~ TreatmentGroup + PoliticalInterest + ReadNews + Education, data = tv)

tidy(reg3)
```

```
##                term     estimate   std.error  statistic      p.value
## 1       (Intercept) -0.144715179 0.085775678 -1.6871354 9.220479e-02
## 2    TreatmentGroup  0.406446411 0.034296486 11.8509638 1.045858e-28
## 3 PoliticalInterest  0.036886861 0.023493045  1.5701183 1.170240e-01
## 4          ReadNews  0.015588089 0.009254475  1.6843841 9.273530e-02
## 5         Education -0.002065028 0.006345252 -0.3254446 7.449816e-01
```

```r
nobs (reg3)
```

```
## [1] 502
```

In this case the assignment to watch the program should be random. Using this as an instrument is useful since it does create a difference in the treatment variable (watching the program), but it does not effect the dependent variable (information levels).

We can test the instrument using an OLS regression with the treatment group and the explanatory variable, and the test confirms that it is a good instrument.

d) 2SLS model

```r
ivreg_tv <- ivreg (InformationLevel ~ WatchProgram + PoliticalInterest + ReadNews + Education | Treatme
                     PoliticalInterest + ReadNews + Education, data = tv)

summary(ivreg_tv)
```

```
##
## Call:
## ivreg(formula = InformationLevel ~ WatchProgram + PoliticalInterest +
##     ReadNews + Education | TreatmentGroup + PoliticalInterest +
##     ReadNews + Education, data = tv)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.5086 -0.5071  0.2001  0.4929  1.9362
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)       1.6887506  0.1646510  10.257  < 2e-16 ***
```

```
## WatchProgram      0.2913412  0.1613956   1.805   0.0717 .
## PoliticalInterest 0.2640640  0.0461014   5.728 1.78e-08 ***
## ReadNews          0.1078766  0.0184163   5.858 8.64e-09 ***
## Education         0.0007689  0.0124353   0.062   0.9507
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7426 on 488 degrees of freedom
## Multiple R-Squared: 0.2043,  Adjusted R-squared: 0.1978
## Wald test:    30 on 4 and 488 DF,  p-value: < 2.2e-16
```

```r
nobs (ivreg_tv)
```

```
## [1] 493
```

From looking at the observations (nobs) we can see that there are 9 more observations in part c. This is because there are some missing variables from information level. Therefore, we have a different result from the first stage of the 2SLS, due to the missing variables.

e) The results of the 2SLS suggest that watching the program and information levels are not corelated at the alpha equals 0.05 level. The results are less significant than part b. We can't be certain that we have defeated endogeneity but using assignment to group as an instrument helps to reduce the endogeneity issue in this experiment. We can see the difference in the number of groups who watched the program and those who did not as the result of the assignment, therefore it is a useful IV.

Note: We obtain different numerical answers even though our processes are exactly the same (even when I replicate the code in the solutions). I am using a .dta file and the solutions use the .csv file, and I am unsure if that is causing the issue.

Question 4:

```r
inmates <- import ("inmates.dta")
```

A) OLS/Linear Model

```r
reg_in <- tidy (lm(prison ~ educ + age + AfAm + factor(state) + factor(year), data = inmates))

head(reg_in)
```

```
##               term      estimate    std.error     statistic      p.value
## 1    (Intercept)  2.911017e-02 4.082012e-04   71.31330309 0.000000e+00
## 2           educ -1.198227e-03 1.391285e-05  -86.12375996 0.000000e+00
## 3            age -3.748158e-04 3.654210e-06 -102.57097019 0.000000e+00
## 4           AfAm  2.117762e-02 1.413494e-04  149.82461390 0.000000e+00
## 5 factor(state)2 -9.472433e-05 1.069555e-03   -0.08856425 9.294282e-01
## 6 factor(state)4  5.170998e-03 5.231901e-04    9.88359148 4.907607e-23
```

From this OLS we can see that education attainment is highly significant in whether or not an individual will go to prison. The coefficient, however, is small but when the difference is between no education and a full 12 years of education, the difference is substantive and the chances of going to prison are much more likely when one has no education (30% probability).

B) No, we cannot prove a causal relationship between education levels and crime. There can be multiple confounding factors at play which can effect the likelihood of crime. This is a difficult relationship to estimate due to the large amount of factors, such as economic conditions, family background, geography and historical discrimination. As a result these confounders could bias the regression.

C) test

```r
test <- lm ( educ ~ age + AfAm + state + year + ca9 +ca10 +ca11, data = inmates)

linearHypothesis(test, c("ca9", "ca10", "ca11"))
```

```
## Linear hypothesis test
##
## Hypothesis:
## ca9 = 0
## ca10 = 0
## ca11 = 0
##
## Model 1: restricted model
## Model 2: educ ~ age + AfAm + state + year + ca9 + ca10 + ca11
##
##    Res.Df      RSS Df Sum of Sq      F    Pr(>F)
## 1 3610662 34559655
## 2 3610659 34464227  3     95428 3332.5 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Since there are multiple IVs we use an f test to discover their strength. The test demonstrates that they are strong instruments.

D) 2sls

```r
ivreg_in <- ivreg (prison ~ educ + age + AfAm + factor(state) + factor(year) | ca9 + ca10 + ca11 + age +
                     AfAm + factor (state) + factor (year), data = inmates)
```

```
## Error: cannot allocate vector of size 1.5 Gb
```

```r
ivreg_in
```

```
## Error in eval(expr, envir, enclos): object 'ivreg_in' not found
```

I am unable to carry out this 2SLS due to the error: cannot allocate vector of size 1.5 GB, even though I have sufficient space on my computer for R to run this. Therefore, I also cannot finish part e.

Note: The code still does not run, but my code matches the solutions.

Question 5:

```r
income <- read_csv ("democracy_income.csv")
```

```
## Parsed with column specification:
## cols(
##   CountryCode = col_integer(),
##   democracy_fh = col_double(),
##   log_gdp = col_double(),
##   year = col_integer(),
##   worldincome = col_double(),
##   YearOrder = col_integer()
## )
```

```r
income <- income %>%
  group_by (CountryCode) %>%
  select (CountryCode, year, democracy_fh, log_gdp, worldincome, YearOrder)
```

a) Pooled Regression

```r
pincome <- pdata.frame(income)

pincome$lag_gdp <- lag(pincome$log_gdp, k = 1)

regplm <- tidy(plm(democracy_fh ~ lag_gdp, data = pincome, model = "pooling"))

regplm
```

```
##          term   estimate   std.error statistic       p.value
## 1 (Intercept) -1.3372658 0.073522000 -18.18865  3.094652e-63
## 2     lag_gdp  0.2337698 0.008984252  26.01995 7.472710e-112
```

From this pooled regression we can see that the lagged GDP is highly significant. But, bias still remains a concern because we cannot be certain that democracy is not influenced by other factors apart from GDP. Alternative theories point to institutional mechanisms, self expression values (Inglehart and Wetzel) or colonial history.

b)

```r
regplm2 <- tidy(plm(democracy_fh ~ lag_gdp + year + CountryCode, data = pincome, model = "pooling"))

head(regplm2)
```

```
##          term     estimate  std.error  statistic      p.value
## 1 (Intercept) -0.193512460 0.21733121 -0.8904035 0.3735263226
## 2     lag_gdp  0.038408436 0.02899967  1.3244437 0.1857471661
## 3    year1965  0.006163702 0.03606057  0.1709264 0.8643263930
## 4    year1970 -0.120493663 0.03620568 -3.3280321 0.0009160006
## 5    year1975 -0.139872667 0.03699592 -3.7807596 0.0001683989
## 6    year1980 -0.089822072 0.03799717 -2.3639150 0.0183293485
```

Here we included country and year fixed effects, there still is a significant result for the lagged GDP (alpha = 0.05). But this significance has reduced once the controls are added.

c) Instrument test

```r
test1 <- tidy(plm (log_gdp ~ worldincome + factor(year) + factor(CountryCode), data = pincome, model =

test1
```

```
##                         term    estimate  std.error  statistic      p.value
## 1                (Intercept)  1.57758874 0.65857094  2.3954728 1.680702e-02
## 2                worldincome  0.40744822 0.05081291  8.0185965 3.395804e-15
## 3          factor(year)1960  0.17750748 0.04446294  3.9922573 7.090488e-05
## 4          factor(year)1965  0.30983410 0.04349223  7.1238948 2.182631e-12
## 5          factor(year)1970  0.37723890 0.04475546  8.4288917 1.414716e-16
## 6          factor(year)1975  0.45815088 0.04562716 10.0411886 1.553826e-22
## 7          factor(year)1980  0.50565090 0.04698569 10.7618063 1.847550e-25
## 8          factor(year)1985  0.51032731 0.04814775 10.5991933 8.710781e-25
## 9          factor(year)1990  0.54485224 0.05016736 10.8606924 7.133413e-26
## 10         factor(year)1995  0.56617193 0.05168923 10.9533824 2.906246e-26
## 11         factor(year)2000  0.65188168 0.05453290 11.9539158 1.260390e-30
## 12    factor(CountryCode)6  5.83583596 0.52605489 11.0935875 7.392421e-27
## 13    factor(CountryCode)8  3.66899347 0.27578271 13.3039286 6.245824e-37
## 14    factor(CountryCode)9  5.01211095 0.35793450 14.0028718 2.279258e-40
## 15   factor(CountryCode)10  2.19501298 0.13420497 16.3556763 1.056850e-52
```

```
## 16   factor(CountryCode)14   2.16361003 0.40377185  5.3584965 1.072234e-07
## 17   factor(CountryCode)16   1.02057520 0.22454234  4.5451347 6.257019e-06
## 18   factor(CountryCode)17   2.33328299 0.40646737  5.7403943 1.299666e-08
## 19   factor(CountryCode)18   3.26744175 0.48191595  6.7801071 2.200072e-11
## 20   factor(CountryCode)24  -4.64144127 0.72973422 -6.3604544 3.229520e-10
## 21   factor(CountryCode)25   2.92437194 0.32510570  8.9951421 1.422912e-18
## 22   factor(CountryCode)26   4.61874651 0.46259966  9.9843275 2.603398e-22
## 23   factor(CountryCode)27   0.45622569 0.21952114  2.0782768 3.797319e-02
## 24   factor(CountryCode)30   4.08647928 0.45482991  8.9846318 1.553154e-18
## 25   factor(CountryCode)31   3.36481128 0.42930050  7.8378928 1.320088e-14
## 26   factor(CountryCode)32   2.68462458 0.13895502 19.3200984 1.332742e-69
## 27   factor(CountryCode)33   1.52175050 0.18935172  8.0366343 2.961190e-15
## 28   factor(CountryCode)34   2.45388155 0.20264430 12.1093047 2.503316e-31
## 29   factor(CountryCode)35   3.20062535 0.44767269  7.1494764 1.830688e-12
## 30   factor(CountryCode)36   0.12329746 0.13769764  0.8954218 3.708065e-01
## 31   factor(CountryCode)37   1.86280838 0.24894420  7.4828350 1.762755e-13
## 32   factor(CountryCode)38  -2.75350014 0.34326134 -8.0215854 3.319670e-15
## 33   factor(CountryCode)39   4.05125693 0.41499387  9.7622091 1.912394e-21
## 34   factor(CountryCode)40   0.49015078 0.15681057  3.1257509 1.831719e-03
## 35   factor(CountryCode)41   1.36639053 0.20072863  6.8071532 1.840945e-11
## 36   factor(CountryCode)42  -0.19437679 0.20788384 -0.9350260 3.500314e-01
## 37   factor(CountryCode)43   5.64499436 0.58375776  9.6700973 4.327502e-21
## 38   factor(CountryCode)44  -1.20210951 0.36096803 -3.3302382 9.038101e-04
## 39   factor(CountryCode)47   3.62561713 0.23055196 15.7258136 2.704172e-49
## 40   factor(CountryCode)51  -3.86362148 0.63075840 -6.1253588 1.362756e-09
## 41   factor(CountryCode)52   2.40567465 0.13405970 17.9448015 1.337022e-61
## 42   factor(CountryCode)53   1.21208274 0.15993325  7.5786787 8.844934e-14
## 43   factor(CountryCode)54   1.96176943 0.18978708 10.3366856 1.025163e-23
## 44   factor(CountryCode)55   2.24849443 0.23842010  9.4308089 3.508712e-20
## 45   factor(CountryCode)56   2.33406838 0.27576447  8.4639925 1.071382e-16
## 46   factor(CountryCode)58   4.32760693 0.33666224 12.8544471 8.819725e-35
## 47   factor(CountryCode)60   2.19494524 0.47992875  4.5734814 5.484141e-06
## 48   factor(CountryCode)61   2.48688384 0.44522556  5.5856717 3.103006e-08
## 49   factor(CountryCode)63   3.50493263 0.21402430 16.3763300 8.148029e-53
## 50   factor(CountryCode)64   0.12530393 0.17891563  0.7003521 4.838926e-01
## 51   factor(CountryCode)65   4.11476207 0.27523499 14.9499961 3.361537e-45
## 52   factor(CountryCode)66  -0.40867343 0.26971455 -1.5152072 1.300790e-01
## 53   factor(CountryCode)67   3.39702591 0.19987560 16.9957009 3.086822e-56
## 54   factor(CountryCode)70   2.75841444 0.41220420  6.6918640 3.918406e-11
## 55   factor(CountryCode)71   2.44724763 0.28354461  8.6309086 2.819818e-17
## 56   factor(CountryCode)72  -3.15629053 0.37604453 -8.3933958 1.872113e-16
## 57   factor(CountryCode)73  -0.07205306 0.19657767 -0.3665373 7.140523e-01
## 58   factor(CountryCode)74  -8.16050381 1.05838839 -7.7103112 3.389366e-14
## 59   factor(CountryCode)75   3.43327632 0.25139798 13.6567377 1.186732e-38
## 60   factor(CountryCode)76  -0.12997308 0.19551334 -0.6647786 5.063663e-01
## 61   factor(CountryCode)77   2.88255482 0.30017642  9.6028690 7.823751e-21
## 62   factor(CountryCode)78  -5.72135013 0.77222104 -7.4089539 2.984374e-13
## 63   factor(CountryCode)79   0.38284769 0.13505991  2.8346510 4.692631e-03
## 64   factor(CountryCode)81  -0.23057148 0.15224368 -1.5144897 1.302608e-01
## 65   factor(CountryCode)82   2.40115591 0.18038988 13.3109240 5.777586e-37
## 66   factor(CountryCode)83   1.22519311 0.19461809  6.2953711 4.833784e-10
## 67   factor(CountryCode)84   4.08733503 0.55095706  7.4186090 2.786635e-13
## 68   factor(CountryCode)85  -2.60482312 0.56982005 -4.5713083 5.539995e-06
## 69   factor(CountryCode)86   4.69106193 0.48581424  9.6560816 4.897450e-21
```

```
## 70   factor(CountryCode)88   2.22364916 0.13409756 16.5823236 6.030435e-54
## 71   factor(CountryCode)89   2.23541498 0.14145772 15.8027074 1.046858e-49
## 72   factor(CountryCode)90   3.85959296 0.25643827 15.0507682 1.003303e-45
## 73   factor(CountryCode)91  -1.29357077 0.29198985 -4.4301909 1.060114e-05
## 74   factor(CountryCode)92  -2.47183194 0.41027604 -6.0248021 2.486561e-09
## 75   factor(CountryCode)93   5.27276878 0.42398482 12.4362207 7.961770e-33
## 76   factor(CountryCode)95   0.65032543 0.19374060  3.3566812 8.226445e-04
## 77   factor(CountryCode)99  -0.62378310 0.31741411 -1.9652028 4.970496e-02
## 78  factor(CountryCode)101   0.10529095 0.17907559  0.5879693 5.567038e-01
## 79  factor(CountryCode)107  -0.27820140 0.21516742 -1.2929532 1.963666e-01
## 80  factor(CountryCode)109  -0.57060193 0.16684338 -3.4199855 6.550438e-04
## 81  factor(CountryCode)110   4.18780881 0.58407716  7.1699582 1.589658e-12
## 82  factor(CountryCode)114   1.57637108 0.19235467  8.1951279 8.791853e-16
## 83  factor(CountryCode)116   2.51806586 0.39434516  6.3854362 2.763713e-10
## 84  factor(CountryCode)118   4.11117724 0.36647963 11.2180238 2.169874e-27
## 85  factor(CountryCode)120   1.65209331 0.31660332  5.2181806 2.254698e-07
## 86  factor(CountryCode)125   1.72423037 0.32430071  5.3167640 1.339925e-07
## 87  factor(CountryCode)126  -1.86372203 0.26810478 -6.9514688 7.040211e-12
## 88  factor(CountryCode)127  -2.41696732 0.49729557 -4.8602229 1.388107e-06
## 89  factor(CountryCode)128  -0.56475844 0.15133105 -3.7319403 2.022055e-04
## 90  factor(CountryCode)129  -4.19758754 0.66615973 -6.3011728 4.663773e-10
## 91  factor(CountryCode)130   5.67950326 0.64333545  8.8282143 5.662703e-18
## 92  factor(CountryCode)131   1.45417353 0.26246958  5.5403507 3.988204e-08
## 93  factor(CountryCode)132   1.13694729 0.24194596  4.6991786 3.030096e-06
## 94  factor(CountryCode)133   0.99873720 0.14194158  7.0362553 3.970314e-12
## 95  factor(CountryCode)134  -1.10309299 0.44038248 -2.5048521 1.243004e-02
## 96  factor(CountryCode)135   2.23809439 0.13407078 16.6933789 1.471401e-54
## 97  factor(CountryCode)136   3.04745728 0.46489181  6.5551967 9.456693e-11
## 98  factor(CountryCode)137   2.96827827 0.16129042 18.4033143 3.068731e-64
## 99  factor(CountryCode)140   2.03094367 0.31644406  6.4180180 2.253841e-10
## 100 factor(CountryCode)141   1.78285727 0.32737288  5.4459529 6.688043e-08
## 101 factor(CountryCode)142  -4.18218590 0.65799444 -6.3559593 3.321126e-10
## 102 factor(CountryCode)144   3.95750701 0.39184539 10.0996647 9.117356e-23
## 103 factor(CountryCode)145   1.50924243 0.18078556  8.3482465 2.669744e-16
## 104 factor(CountryCode)147  -2.38192891 0.40175595 -5.9287956 4.379659e-09
## 105 factor(CountryCode)148   4.56252831 0.43115942 10.5819984 1.025219e-24
## 106 factor(CountryCode)150   1.57963017 0.13406481 11.7825859 7.363100e-30
## 107 factor(CountryCode)151   3.65561499 0.37418804  9.7694597 1.792868e-21
## 108 factor(CountryCode)153   2.60061461 0.28882268  9.0041910 1.319477e-18
## 109 factor(CountryCode)155   2.76096226 0.44537605  6.1991710 8.715796e-10
## 110 factor(CountryCode)160   0.37069777 0.15102102  2.4546104 1.429638e-02
## 111 factor(CountryCode)162 -24.85640410 3.32420773 -7.4773919 1.832753e-13
## 112 factor(CountryCode)165   2.52657171 0.37846271  6.6758802 4.347218e-11
## 113 factor(CountryCode)166   1.82240260 0.17410140 10.4674780 3.018992e-24
## 114 factor(CountryCode)168  -2.66790570 0.31767422 -8.3982443 1.801933e-16
## 115 factor(CountryCode)172   2.47369802 0.13521491 18.2945650 1.304973e-63
## 116 factor(CountryCode)174   0.47499776 0.20108471  2.3621774 1.838478e-02
## 117 factor(CountryCode)175   3.15920206 0.36165368  8.7354346 1.209123e-17
## 118 factor(CountryCode)176   3.33618147 0.48422523  6.8897307 1.064373e-11
## 119 factor(CountryCode)177  -2.19831011 0.26530397 -8.2860055 4.343401e-16
## 120 factor(CountryCode)178   0.93909725 0.14527962  6.4640674 1.686853e-10
## 121 factor(CountryCode)182  -0.17887470 0.26040919 -0.6868986 4.923276e-01
## 122 factor(CountryCode)183   0.12968913 0.16743508  0.7745637 4.388056e-01
## 123 factor(CountryCode)184   3.84136083 0.37846368 10.1498797 5.757256e-23
```

```
## 124 factor(CountryCode)187   2.04134043 0.42286300  4.8274274 1.630333e-06
## 125 factor(CountryCode)188   2.39670020 0.44157433  5.4276257 7.387499e-08
## 126 factor(CountryCode)191   3.35676657 0.26603560 12.6177347 1.142064e-33
## 127 factor(CountryCode)192   6.39153415 0.49743132 12.8490787 9.350476e-35
## 128 factor(CountryCode)195  -2.38782601 0.44652965 -5.3475195 1.137137e-07
## 129 factor(CountryCode)196   3.22198920 0.23902336 13.4798087 8.733196e-38
## 130 factor(CountryCode)197   4.81742382 0.67751828  7.1103968 2.394308e-12
## 131 factor(CountryCode)208   3.52451096 0.28789781 12.2422291 6.208334e-32
## 132 factor(CountryCode)209   2.70431851 0.45369222  5.9606896 3.632056e-09
## 133 factor(CountryCode)210  -1.02871838 0.15915131 -6.4637756 1.689963e-10
## 134 factor(CountryCode)211   2.83198231 0.34082547  8.3091861 3.624641e-16
```

The instrument test shows that world income is a strong instrument for log gdp. 1. Inclusion condition: The test determines this, that world income is corelated with log gdp at the alpha = 0.05 level. 2. Exclusion condition: The IV should not be corelated with y, only through x. This can only be justified through theory, and here trading partner income is corelated with the GDP of a country, but should not be corelated with the level of democracy.

d)

```r
reg <- tidy (plm(democracy_fh ~ lag_gdp + year + CountryCode | lag(worldincome, k = 1) + year + Country

head(reg)
```

```
##          term    estimate  std.error  statistic    p.value
## 1 (Intercept)  1.51163356 0.79580909  1.8994927 0.05787170
## 2     lag_gdp -0.21298773 0.11645759 -1.8288866 0.06780174
## 3    year1965  0.05206503 0.04299022  1.2110902 0.22623096
## 4    year1970 -0.03297144 0.05450022 -0.6049781 0.54537111
## 5    year1975 -0.01995469 0.06616449 -0.3015921 0.76304415
## 6    year1980  0.05824546 0.07725812  0.7539073 0.45113444
```

With the use of world income as an instrument we find that the coefficient becomes negative. Therefore, there is a negative relationship between GDP and democracy in this model. This has the opposite outcome from the OLS model and the panel data which demonstrate a positive relationship between democracy and GDP. Also the significance of the result disappears, thus we cannot reject the null hypothesis.