```
library(stargazer)
library(dplyr)
library(haven)
library(AER)
library(sandwich)

pums80 = read_dta("C:/Users/Bernardo/Documents/GitHub/angrist-evans/data/pums80.dta")

# Table 1

summ_stats = pums80 %>%
  filter(agem1 >= 19 & 35 >= agem1) %>%
  summarise('Mean children ever born' = mean(kidcount),
            'Percent with more than 2 children' = 100 * length(which(pums80$kidcount > 2))/nrow(pums80),
            'Percent worked last year' = 100 * length(which(pums80$workedm == 1))/nrow(pums80),
            'Observations' = nrow(pums80))

summ_stats = as.data.frame(t(summ_stats))
names(summ_stats)[1] <- "1980 PUMS"
stargazer(summ_stats, type = "html", summary = FALSE,  out="tab1.htm", digits = 2,
          title = "TABLE 1-FERTILITY AND LABOR-SUPPLY MEASURES")

# Creating Dummies

pums80 = pums80 %>%
  mutate(twoboys = ifelse(boy1st == 1 & boy2nd == 1, 1, 0)) %>%
  mutate(twogirls = ifelse(boy1st == 0 & boy2nd == 0, 1, 0)) %>%
  mutate(morekids = ifelse(kidcount > 2, 1, 0))

# Table 2

stargazer(as.data.frame(pums80[c("kidcount", "morekids", "boy1st", "boy2nd", "twoboys", "twogirls", "samesex",
                   "agem1", "agefstm", "workedm", "weeksm1", "hourswm", "incomem")]),
          covariate.labels=c("Children ever born",
                             "More than 2 children (=1 if mother had more than 2 children, =0 otherwise)",
                             "Boy 1st (=1 if first child was a boy)",
                             "Boy 2nd (=1 if second child was a boy)",
                             "Two boys (=1 if first two children were boys)",
                             "Two girls (=1 if first two children were girls)",
                             "Same sex (=1 if first two children were the same sex)",
                             "Age",
                             "Age at first birth (parent's age in years when first child was born)",
                             "Worked for pay (=1 if worked for pay in year prior to census)",
                             "Weeks worked (weeks worked in year prior to census)",
                             "Hours/week (average hours worked per week)",
                             "Labor income (labor earnings in year prior to census, in 1995 dollars)"),
          type = "html",  summary.stat = c("mean", "sd"),
          title = "TABLE2-DESCRIPTIVE STATISTICS, WOMEN AGED 21-35 WITH 2 OR MORE CHILDREN", out="tab2.htm")

# Models for Table 3
# worked for pay

wfp_ols = lm(data = pums80, formula = workedm ~ morekids + agem1 + agefstm + black + hispan + othrace +
               boy1st + boy2nd)
cov1 = vcovHC(wfp_ols, type = "HC1")
robust_se1 = sqrt(diag(cov1))

wfp_reduced = lm(data = pums80, formula = workedm ~ agem1 + agefstm + black + hispan + othrace +
                   boy1st + boy2nd + samesex)

wfp_1stage = lm(morekids ~ agem1 + agefstm + black + hispan + othrace +
                    boy1st + boy2nd + samesex, data = pums80)
pums80$morekids_hat1 = predict(wfp_1stage, pums80)
wfp_2stage = lm(workedm ~ morekids_hat1 + agem1 + agefstm + black + hispan + othrace +
                  boy1st + boy2nd + samesex, data = pums80)
cov3 = vcovHC(wfp_2stage, type = "HC1")
robust_se3 = sqrt(diag(cov3))

wfp_2sls = ivreg(workedm ~ morekids + agem1 + agefstm + black + hispan + othrace +
               boy1st + boy2nd | . - morekids + samesex, data = pums80)
cov4 = vcovHC(wfp_2sls, type = "HC1")
robust_se4 = sqrt(diag(cov4))

CIV1 = summary(wfp_reduced)$coefficients[9]/summary(wfp_1stage)$coefficients[9]

wfp_1 = lm(workedm ~ 1, data = pums80)
cov2 = vcovHC(wfp_1, type = "HC1")
robust_se2 = sqrt(diag(cov2))

# weeks worked

ww_ols = lm(data = pums80, formula = weeksm1 ~ morekids + agem1 + agefstm + black + hispan + othrace +
               boy1st + boy2nd)
cov5 = vcovHC(ww_ols, type = "HC1")
robust_se5 = sqrt(diag(cov5))
```

```
ww_reduced = lm(data = pums80, formula = weeksm1 ~ agem1 + agefstm + black + hispan + othrace +
                  boy1st + boy2nd + samesex)

ww_2stage = lm(weeksm1 ~ morekids_hat1 + agem1 + agefstm + black + hispan + othrace +
                 boy1st + boy2nd + samesex, data = pums80)
cov7 = vcovHC(ww_2stage, type = "HC1")
robust_se7 = sqrt(diag(cov7))

ww_2sls = ivreg(weeksm1 ~ morekids + agem1 + agefstm + black + hispan + othrace +
                  boy1st + boy2nd | . - morekids + samesex, data = pums80)
cov8 = vcovHC(ww_2sls, type = "HC1")
robust_se8 = sqrt(diag(cov8))

CIV2 = summary(ww_reduced)$coefficients[9]/summary(wfp_1stage)$coefficients[9]

ww_1 = lm(weeksm1 ~ 1, data = pums80)
cov6 = vcovHC(ww_1, type = "HC1")
robust_se6 = sqrt(diag(cov6))

stargazer(wfp_ols, wfp_1, wfp_2stage, wfp_2sls, ww_ols, ww_1, ww_2stage, ww_2sls, type = "html",
          covariate.labels=c("More than 2 children",
                             "More than 2 children",
                             "Age",
                             "Age at first birth",
                             "Black",
                             "Hispanic",
                             "Other race",
                             "Boy 1st",
                             "Boy 2nd",
                             "Same sex",
                             "Constant"), style = "qje",
          dep.var.labels   = c("Worked for pay","Weeks worked"),
          column.labels = c("OLS", "Cov Adjusted", "Manual 2SLS", "IVReg", "OLS", "Cov Adjusted", "Manual 2SLS",
"IVReg"),
          out="tab3.htm", se = list(robust_se1, robust_se2, robust_se3, robust_se4, robust_se5, robust_se6,
robust_se7, robust_se8))

# Models for Table 4
# hours worked

hours_ols = lm(data = pums80, formula = hourswm ~ morekids + agem1 + agefstm + black + hispan + othrace +
                 boy1st + boy2nd)
cov1 = vcovHC(hours_ols, type = "HC1")
robust_se1 = sqrt(diag(cov1))

hours_reduced = lm(data = pums80, formula = hourswm ~ agem1 + agefstm + black + hispan + othrace +
                     boy1st + boy2nd + samesex)

hours_1stage = lm(morekids ~ agem1 + agefstm + black + hispan + othrace +
                    boy1st + boy2nd + samesex, data = pums80)
pums80$morekids_hat1 = predict(hours_1stage, pums80)
hours_2stage = lm(hourswm ~ morekids_hat1 + agem1 + agefstm + black + hispan + othrace +
                    boy1st + boy2nd + samesex, data = pums80)
cov3 = vcovHC(hours_2stage, type = "HC1")
robust_se3 = sqrt(diag(cov3))

hours_2sls = ivreg(hourswm ~ morekids + agem1 + agefstm + black + hispan + othrace +
                     boy1st + boy2nd | . - morekids + samesex, data = pums80)
cov4 = vcovHC(hours_2sls, type = "HC1")
robust_se4 = sqrt(diag(cov4))

CIV1 = summary(hours_reduced)$coefficients[9]/summary(hours_1stage)$coefficients[9]

hours_1 = lm(hourswm ~ 1, data = pums80)
cov2 = vcovHC(hours_1, type = "HC1")
robust_se2 = sqrt(diag(cov2))

# labor income

linc_ols = lm(data = pums80, formula = incomem ~ morekids + agem1 + agefstm + black + hispan + othrace +
                boy1st + boy2nd)
cov5 = vcovHC(linc_ols, type = "HC1")
robust_se5 = sqrt(diag(cov5))

linc_reduced = lm(data = pums80, formula = incomem ~ agem1 + agefstm + black + hispan + othrace +
                    boy1st + boy2nd + samesex)

linc_2stage = lm(incomem ~ morekids_hat1 + agem1 + agefstm + black + hispan + othrace +
                   boy1st + boy2nd + samesex, data = pums80)
cov7 = vcovHC(linc_2stage, type = "HC1")
robust_se7 = sqrt(diag(cov7))

linc_2sls = ivreg(incomem ~ morekids + agem1 + agefstm + black + hispan + othrace +
                    boy1st + boy2nd | . - morekids + samesex, data = pums80)
```

```
cov8 = vcovHC(linc_2sls, type = "HC1")
robust_se8 = sqrt(diag(cov8))

CIV2 = summary(linc_reduced)$coefficients[9]/summary(hours_1stage)$coefficients[9]

linc_1 = lm(incomem ~ 1, data = pums80)
cov6 = vcovHC(linc_1, type = "HC1")
robust_se6 = sqrt(diag(cov6))

stargazer(hours_ols, hours_1, hours_2stage, hours_2sls, linc_ols, linc_1, linc_2stage, linc_2sls, type = "html",
          covariate.labels=c("More than 2 children",
                             "More than 2 children",
                             "Age",
                             "Age at first birth",
                             "Black",
                             "Hispanic",
                             "Other race",
                             "Boy 1st",
                             "Boy 2nd",
                             "Same sex",
                             "Constant"), style = "qje",
          dep.var.labels   = c("Hours worked per week","Labor income"),
          column.labels = c("OLS", "Cov Adjusted", "Manual 2SLS", "IVReg", "OLS", "Cov Adjusted", "Manual 2SLS",
"IVReg"),
          out="tab4.htm", se = list(robust_se1, robust_se2, robust_se3, robust_se4, robust_se5, robust_se6,
robust_se7, robust_se8))
```