Homework Assignment 5

Deadline: April 17, midnight

Source: Stock and Watson, 4th Edition, Exercise 7.1

Data description: You can find the data description here.

Questions

a. Regress (i) `birthweight` on `smoker`, `alcohol`,`nprevist`, and `unmarried`. Interpret the coefficient on `unmarried`.

b. Construct a 95% confidence interval on for the coefficient. Is it statistically significant? Is the magnitude of the coefficient large?

c. Looking at this regression, a family advocacy group claims that higher rates of marriage will lead to healthier babies thus one obviuous public policy is to encourage marriage. Do you agree?

d. Consider the data set that you have and briefly discuss what variables can be added to the regression to help to solve question (c).

e. Run the regression with these additional controls. How did the coefficient on marriage has changed with these additional controls.

Header for the R script

Start a new R script, copy/paste the header below and save it to `Dropbox\EC282\Assignment4` or a similar path that you created for this homework assignment. Run the R script and make sure that you have the data `df1` in your environment. Conduct the analysis below the header.

```
################################################################################
# list the packages we need and loads them, installs them automatically if we don't have them
# add any package that you need to the list
need <- c('glue', 'dplyr','readxl', 'ggplot2','tidyr','AER','scales','mvtnorm',
          'stargazer','httr', 'repmis')

have <- need %in% rownames(installed.packages())
if(any(!have)) install.packages(need[!have])
invisible(lapply(need, library, character.only=T))

# Save the R script to the assignment 1 folder before this
# To set up the working directory
getwd()
setwd(getwd()) #change getwd() here is you need to set a different working directory


#this clears the workspace
rm(list = ls())
#this sets the random number generator seed to my birthday for replication

options(scipen=999)
################################################################################
#get the data url
df1.url <- 'https://www.dropbox.com/s/z8r6hc0r4ytt4f8/birthweight_smoking.xlsx?dl=1'
#download the data
GET(df1.url, write_disk(tdf <- tempfile(fileext = ".xlsx")))
#check if it worked
```

```
df1 <- read_excel(tdf) %>%
    mutate(birthweight = birthweight + rnorm(length(birthweight)) * 50)



head(df1)

#CONDUCT THE ANALYSIS BELOW
```