

## PS2\_part2

```
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(forcats)
library(broom)
library(modelr)

##
## Attaching package: 'modelr'

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

library(stringr)
library(lmtest)

## Loading required package: zoo

##
## Attaching package: 'zoo'

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

library(car)

##
## Attaching package: 'car'

## The following object is masked from 'package:dplyr':
##
##   recode

## The following object is masked from 'package:purrr':
##
##   some

library(RColorBrewer)

#options(digits = 3)
#set.seed(1234)
#theme_set(theme_minimal())
#setwd("~/Google Drive/spring2017/mac30200/MACS30200proj/ProblemSets/PS3/")
```

```
biden_df = read.csv("data/biden.csv") %>%
  na.omit() %>%
  mutate(dem = factor(dem),
         rep = factor(rep))
```

(a)

```
biden_interaction = lm(biden ~ age*educ, data = biden_df)
tidy(biden_interaction)

##           term      estimate std.error statistic      p.value
## 1 (Intercept) 38.3735103  9.56356681   4.012468 6.254443e-05
## 2          age  0.6718750  0.17049152   3.940812 8.430505e-05
## 3          educ  1.6574253  0.71399213   2.321350 2.037897e-02
## 4    age:educ -0.0480341  0.01290186  -3.723037 2.028851e-04

instant_effect <- function(model, mod_var){
  int.name <- names(model$coefficients)[[which(str_detect(names(model$coefficients), ":"))]]

  marg_var <- str_split(int.name, ":")[[1]][[which(str_split(int.name, ":")[[1]] != mod_var)]]

  beta.hat <- coef(model)
  cov <- vcov(model)

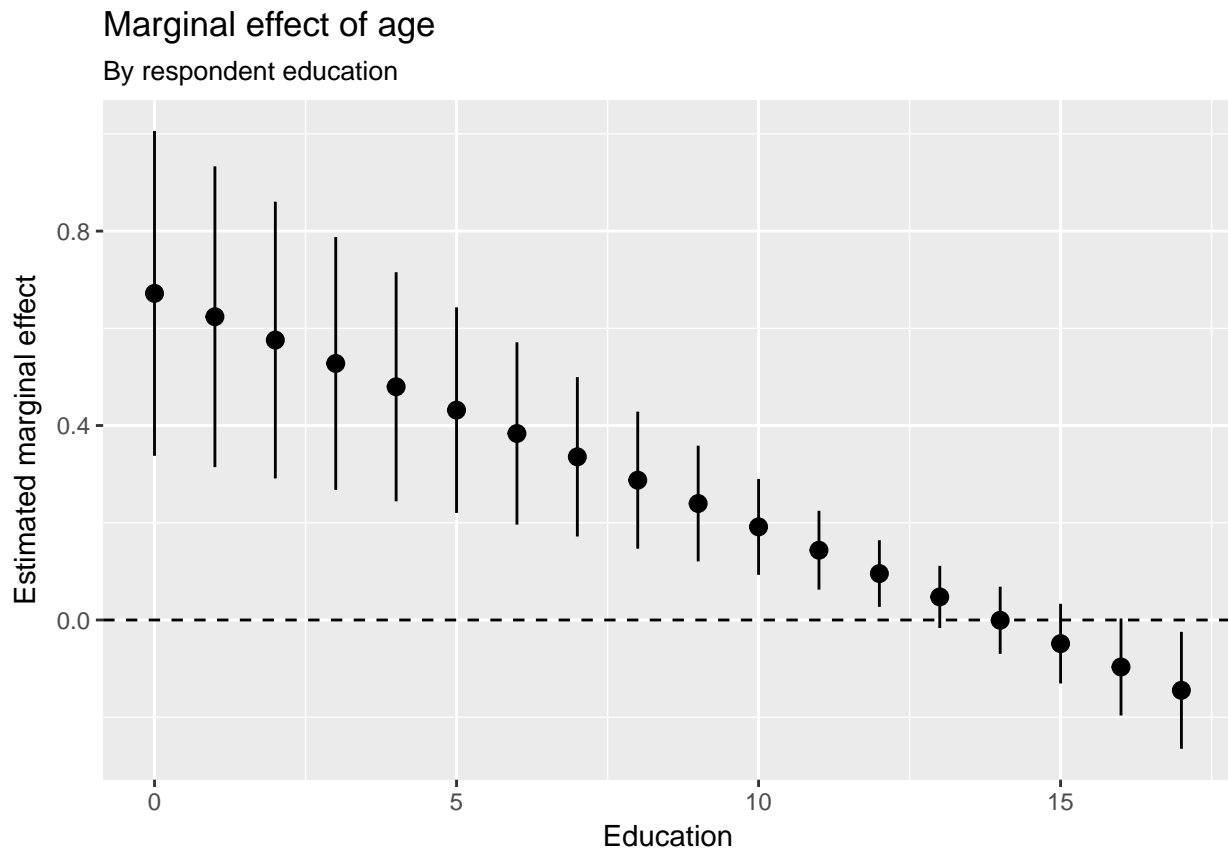
  if(class(model)[1] == "lm"){
    z <- seq(min(model$model[[mod_var]]), max(model$model[[mod_var]]))
  }else{
    z <- seq(min(model$data[[mod_var]]), max(model$data[[mod_var]]))
  }

  dy.dx <- beta.hat[[marg_var]] + beta.hat[[int.name]] * z

  se.dy.dx <- sqrt(cov[marg_var, marg_var]+
                   z^2 * cov[int.name, int.name] +
                   2 * z * cov[marg_var, int.name])

  data_frame(z = z,
             dy.dx = dy.dx,
             se = se.dy.dx)
}

instant_effect(biden_interaction, "educ") %>%
  ggplot(aes(z, dy.dx,
             ymin = dy.dx - 1.96 * se,
             ymax = dy.dx + 1.96 * se)) +
  geom_pointrange() +
  geom_hline(yintercept = 0, linetype = 2) +
  labs(title = "Marginal effect of age",
       subtitle = "By respondent education",
       x = "Education",
       y = "Estimated marginal effect")
```



The marginal effect of age is estimated above. According to the marginal effect plot, marginal effect of age is positive if the education is less than 14 years and is negative if the education is more than 14 years.

```
linearHypothesis(biden_interaction, "age + age:educ")
```

```
## Linear hypothesis test
##
## Hypothesis:
## age + age:educ = 0
##
## Model 1: restricted model
## Model 2: biden ~ age * educ
##
##   Res.Df    RSS Df Sum of Sq    F    Pr(>F)
## 1   1804 985149
## 2   1803 976688  1    8461.2 15.62 8.043e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The test indicates that the marginal effect of age is significant.

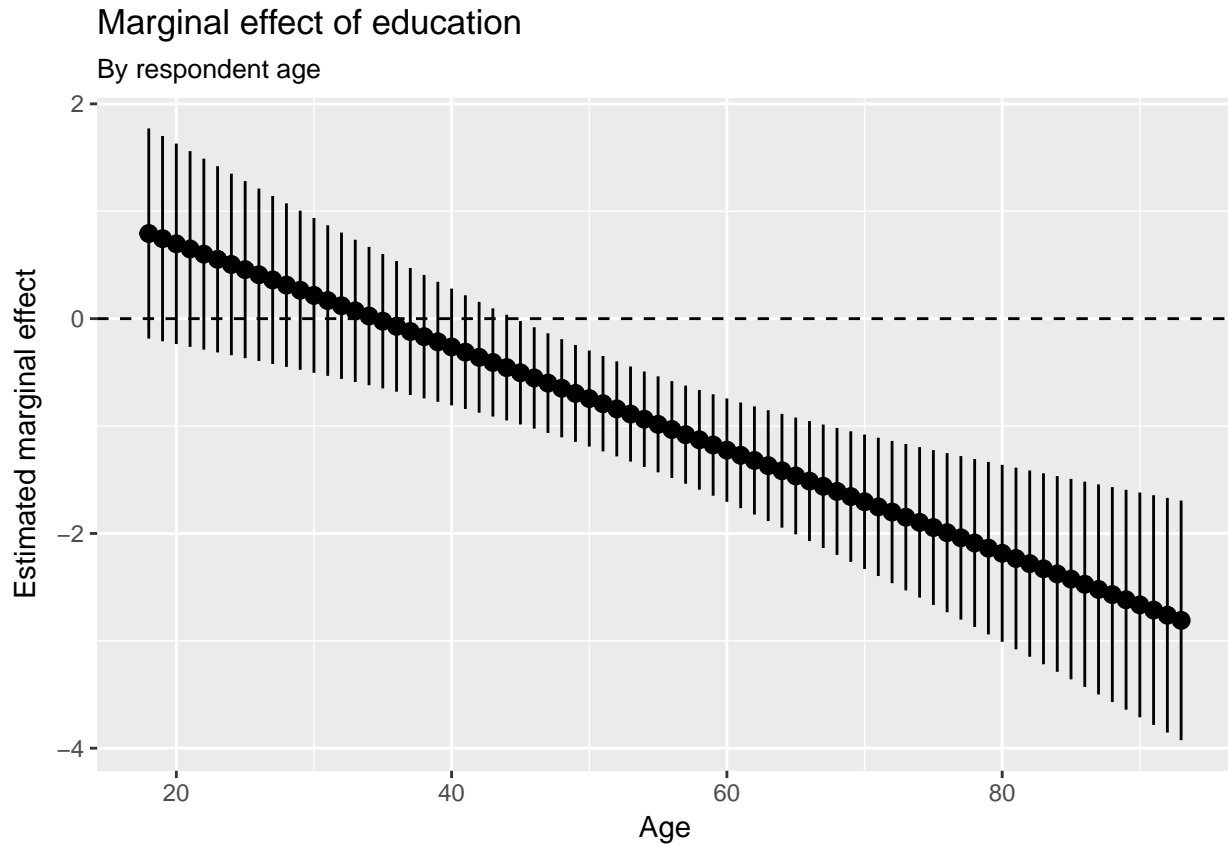
(b)

```
instant_effect(biden_interaction, "age") %>%
  ggplot(aes(z, dy.dx,
             ymin = dy.dx - 1.96 * se,
```

```

      ymax = dy.dx + 1.96 * se)) +
geom_pointrange() +
geom_hline(yintercept = 0, linetype = 2) +
labs(title = "Marginal effect of education",
      subtitle = "By respondent age",
      x = "Age",
      y = "Estimated marginal effect")

```



The marginal effect of education is estimated above. According to the marginal effect plot, marginal effect of education is positive if the age is less than 34 and is negative if the age is more than 34 years.

```
linearHypothesis(biden_interaction, "educ + age:educ")
```

```

## Linear hypothesis test
##
## Hypothesis:
## educ + age:educ = 0
##
## Model 1: restricted model
## Model 2: biden ~ age * educ
##
##   Res.Df    RSS Df Sum of Sq    F Pr(>F)
## 1    1804 979537
## 2    1803 976688  1    2849.1 5.2595 0.02194 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

The test indicates that the marginal effect of education is significant.