

Likelihood to have bad habits and age

[28287667]—6/25/2015

Introduction

As more people start quitting their bad habits, age has played a crucial role. Intuitive, my prediction is that older people are less likely to smoke. Thus, I want to examine the ceteris paribus achievement effects of increasing in age, and estimate the size of the effect.

Review of past research

Self-care in Later Life: Research, Program, and Policy Issues by Marcia G. Ory (1998) talks about the healthy and bad habits with respect to age. Ory has discovered numerous statistics, for example, only 8 percent of people who are over 85 years old who used to have bad habits still have them compare to 34 percent for the ones in their 40s. Thus, Ory concludes that older people in general are more likely than the younger ones to have good habits or they have quitted their bad habits when they were young.

Data

In this report, I explore the relationship between likelihood to have bad habits and age using data from Stock and Watson's *Introduction to Econometrics*. (Stock)

Introduction to Econometrics by Stock and Watson is a book that has gathered cross-sectional data on over 10,000 observations. The data was collected from National Health Interview Survey in 1991 and 1993 from diversified population in terms of smoking, education levels, races, sex, and age.

For this report, I focus on the discovering the relationship between age and likelihood to have bad habits. In the meantime, I observe the education levels of people, and whether they are smokers as the covariates. Table 1 below presents the summary statistics of variables used in the report.

Table 1: Summary statistics						
Variable name	Description	N	Mean	SD	Min	Max
Smoker	If the person smokes=1, if otherwise =0	10000	.2423	.43	0	1
Age	Age in years	10000	28.69	12.11	18	88
hsdrop	If high school dropout=1, if otherwise =0	10000	.091	.29	0	1
Hsgrad	If high school graduate=1, if otherwise =0	10000	.33	.47	0	1
Colsome	If some college=1, if otherwise=0	10000	.28	.45	0	1
Colgrad	If college graduate=1, if otherwise=0	10000	.20	.40	0	1
Age2	Age squared in years	10000	1643.89	1028.75	324	7744

Econometric model

My econometric model 1 and model 2 are multivariate regressions of the forms where \mathbf{x}_i is a vector of control variables.

$$\text{smoker}_i = \beta_0 + \beta_1 \text{age}_i + \boldsymbol{\gamma} \cdot \mathbf{x}_i + \epsilon_i \quad (1)$$

$$\text{smoker}_i = \beta_0 + \beta_1 \text{age}_i + \beta_2 \text{age}_i^2 + \boldsymbol{\gamma} \cdot \mathbf{x}_i + \epsilon_i \quad (2)$$

My hypothesis that the likelihood to smoke decreases as age increases for both models. And my hypothesis would be : $H_0: \beta_{1,2} = 0$, $H_a: \beta_1 > 0$ or $\beta_2 > 0$. I used ordinary least squares to estimate the parameters of the model 1. I also tried using quadratic linear probability model because by a scatter plot I did on the data, it shows more like a parabola instead of linear. Thus, I generated age squared for my model 2. Here are my results:

$$\text{smoker}_i = .1422 + (-.0012)\text{age}_i + \boldsymbol{\gamma} \cdot \mathbf{x}_i + \epsilon_i$$

$$\text{smoker}_i = -.0600 + .0091\text{age}_i - .0001\text{age}_i^2 + \boldsymbol{\gamma} \cdot \mathbf{x}_i + \epsilon_i$$

Results

Table 2 on the right displays coefficient estimates and standard errors for the parameters of equations. According to the table, I find that age is significantly related to the likelihood to smoke. For instance, according to the 95% confidence interval, for every one year on age increases, the likelihood to smoke decreases by .0012 for model 1. For model 2, for the same confidence interval, one more year old increases and the likelihood to smoke decreases by an amount that cannot be expressed easily since age is squared. Overall, the adjusted R squared for model 2 is .0462 which only increases very little from R squared in mode 1 which is .0441. Thus, both models have fitted my prediction and there is no better one between the two.

Conclusion

This report explored the relationship between age and likelihood to have bad habits. And I find that there is significant relationship between them. As age increases, the likelihood to have bad habits decreases.

Table 2: Regression results

Variables	Model 1 (smoker)	Model 2 (smoker)
age	-.0012 (.0003)	.0091 (.0020)
age2	/	-.0001 (0)
hsdrop	.2814 (.0183)	.2934 (.0191)
Hsgrad	.2180 (.0121)	.2261 (.0150)
Colsome	.1471 (.0123)	.1547 (.0153)
Colgrad	.0383 (.0119)	.0413 (.0161)
R^2	.0441	.0468
N	10000	10000

Table 1: OLS estimates for equation (1). Standard errors are in parentheses, and are clustered by school.

References

Stock, J., & Watson, M. (2007). Student Resources. Retrieved June 20, 2015, from http://wps.aw.com/aw_stock_ie_2/0,12040,3332253,00.html

Ory, M. (1998). Self-care in later life research, program, and policy issues. New York: Springer Pub.

[Ran Jin]