

Exercise 1

Use the CPS (cps09mar) dataset posted at  
- <https://www.ssc.wisc.edu/~bhansen/econometrics/>

The data is posted in text, excel, and stata formats.

A data description is posted and is included here.

The recommended software is R.

1. Extract the sample of Hispanic women ( $n=3003$ ).
2. Create the wage variable as annual earnings divided by number of weeks worked and average number of hours worked per week.
3. Create the dependent variable by taking the natural logarithm.
4. Create the experience variable as age less years of education, less 6. Divide by its maximum variable in the sample to standardize between 0 and 1.  
(Otherwise the polynomial regression will become numerically unstable.)
5. Create a dummy variable for *married*, three regions, and having attended *college* (education is 16 years or higher).
6. Create dummy variables for education levels equaling 12, 13, 14, 16, 18, and 20.
7. Estimate a wage regression. Regress  $\log(\text{wage})$  on *college*, *experience*, its square, *married*, and the three regions.
8. Calculate the BIC, AIC, and CV criteria.
9. Repeat with a regression replacing *college* with *education* (a linear specification).
10. Repeat with a regression replacing *education* with the six education dummies.
11. Repeat adding a cubic in experience.
12. Repeat adding a 3<sup>rd</sup> and 4<sup>th</sup> power in experience.
13. Repeat adding a 3<sup>rd</sup>, 4<sup>th</sup>, and 5<sup>th</sup> power in experience.
14. Repeat adding a 3<sup>rd</sup>, 4<sup>th</sup>, and 5<sup>th</sup>, and 6<sup>th</sup> power in experience.
15. Create a table of the criteria to make an easy comparison.
16. What do you infer about which model(s) are appropriate for this wage regression?  
Based on this evidence what is your preferred model?

This document describes the dataset cps09mar

The Current Population Survey (CPS) is a monthly survey of about 57,000 U.S. households conducted by the Bureau of the Census of the Bureau of Labor Statistics. The CPS is the primary source of information on the labor force characteristics of the U.S. population. The survey covers employment, earnings, educational attainment, income, poverty, health insurance coverage, job experience, voting and registration, computer usage, veteran status, and other variables. Details can be found at [www.census.gov/cps](http://www.census.gov/cps) and [dataferrett.census.gov](http://dataferrett.census.gov).

From the March 2009 survey we extracted the individuals with non-allocated variables who were full-time employed (defined as those who had worked at least 36 hours per week for at least 48 weeks the past year), and excluded those in the military. This sample has 50,742 individuals. We extracted 12 variables from the CPS on these individuals.

The data are presented in Excel, Stata, and text file formats.

The variables are:

1. age: years, capped at 85
2. female: 1 if female, 0 otherwise
3. hisp: 1 if Spanish, Hispanic, or Latino, 0 otherwise
4. education
  - 0 Less than 1<sup>st</sup> grade
  - 4 1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup>, or 4<sup>th</sup> grade
  - 6 5<sup>th</sup> or 6<sup>th</sup> grade
  - 8 7<sup>th</sup> or 8<sup>th</sup> grade
  - 9 9<sup>th</sup> grade
  - 10 10<sup>th</sup> grade
  - 11 11<sup>th</sup> grade or 12<sup>th</sup> grade with no high school diploma
  - 12 High school graduate, high school diploma or equivalent
  - 13 Some college but no degree
  - 14 Associate degree in college, including occupation/vocation programs
  - 16 Bachelor's degree or equivalent (BA, AB, BS)
  - 18 Master's degree (MA, MS, MENG, MED, MSW, MBA)
  - 20 Professional degree or Doctorate degree (MD, DDS, DVM, LLB, JD, PHD, EDD)
5. earnings: total annual wage and salary earnings
6. hours: number of hours worked per week
7. week: number of weeks worked per year
8. union: 1 for member of a labor union, 0 otherwise
9. uncov: 1 if covered by a union or employee association contract, 0 otherwise
10. region
  - 1 Northeast
  - 2 Midwest
  - 3 South
  - 4 West
11. race
  - 1 White only
  - 2 Black only
  - 3 American Indian, Alaskan Native (AI) only

- 4 Asian only
- 5 Hawaiian/Pacific Islander (HP) only
- 6 White-Black
- 7 White-AI
- 8 White-Asian
- 9 White-HP
- 10 Black-AI
- 11 Black-Asian
- 12 Black-HP
- 13 AI-Asian
- 14 Asian-HP
- 15 White-Black-AI
- 16 White-Black-Asian
- 17 White-AI-Asian
- 18 White-Asian-HP
- 19 White-Black-AI-Asian
- 20 2 or 3 races
- 21 4 or 5 races
- 12. marital
  - 1 Married - civilian spouse present
  - 2 Married - Armed Forces spouse present
  - 3 Married - spouse absent (except separated)
  - 4 Widowed
  - 5 Divorced
  - 6 Separated
  - 7 Never married