CPS code on gender gap

Evan Soltas <esoltas@princeton.edu> Thu, Jan 14, 2016 at 1:04 PM
To: david.backus@gmail.com

Hi Dave,

Happy to help. I'm afraid it would take me quite a while to track down the code for that post, but I can explain how to do this in general. It's pretty straightforward, unless you've never used IPUMS/CPS data before (let me know if that is the case).

1. Go the IPUMS website (https://cps.ipums.org/cps/) and specify the sample as the 2014 ASEC (the March supplement to the CPS).

2. Obviously, you need the gender variable and the hourly wage variable. There is some minor cleaning required for both variables -- I think you have to set up the gender as a binary and remove the not-specifieds, and I think there might be something similar for removing people who are top-coded in the wage distribution (which shows up as a field of "99" or something like that). Then you do the natural log on the wage.

3. In Stata, this is a one-line entry to produce the graph:

tw kdensity lwage if female == 0 | kdensity lwage if female == 1

I am learning Python right now and am a very basic user, so unfortunately I do not know how to do kernel density plots in Python. But I am sure there is a package somewhere that makes that easy.

Best,

Evan
[Quoted text hidden]