Econometrics

Exercise: Frisch-Waugh-Lovell Theorem

In this short exercise you will 'show' that the Frisch-Waugh-Lovell Theorem in fact works.

Frisch-Waugh-Lovell Theorem

Say you have a regression you wish to estimate: $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_k X_k$

If you run OLS on the above you get $\hat{\beta}_1$ for the coefficient on X_1

The theorem says there is also another way to get the exact same answer:

1. Regress X_1 on $X_2, X_3, ..., X_k$, and denote $\tilde{X}_1$ the residuals.
2. Regress Y on $X_2, X_3, ..., X_k$, and denote $\tilde{Y}$ the residuals.
3. Regress $\tilde{Y}$ on $\tilde{X}_1$

The coefficient on $\tilde{X}_1$ is the same as the coefficient on X_1 from a normal OLS regression (all regressions include an intercept).

Think about this and how we interpret $\hat{\beta}$. This tells us how much Y changes when X_1 changes *holding the other Xs constant.*

Well step 1 gives us how X_1 moves when the other Xs are constant (this is the residual), and step 2 tells us how Y moves when the other Xs are constant (we remove that association and use the residuals). Then regress what's left on each other.

TURN IN

Use the data set 'cps_2009.csv'

Let $Y = wage$, $X_1 = age$, $X_2 = sex$, $X_3 = ed$.

Show that the FWL Theorem works using X_1 to show you can recover the coefficient both ways.

Note that if you run a regression and call it out1 as this (NOTE: this is just an example, you will have to write it so that it matches your data and what you actually want to do):

out1 <- lm(y∼x1+x2,data = mydata)

Then the residuals are stored as 'out1$residuals' and can be used just like a vector data in your data set.

Do this in a script file. Make sure it works and you get results.
1. Copy your script file into a word file 2. Copy your results from both outputs in the same file

---