651 Empirical Economics: Assignment 2

Per Hjertstrand*　　Andrew Proctor†

September 2017

8.1 Statements (ii)-(iii) are true, but not (i).

8.2 Let $w_i = inc^2$, then multiplying each term by $\sqrt{w_i} = inc$, we get:

$$\frac{beer}{inc} = \beta_0 \frac{1}{inc} + \beta_1 + \beta_2 \frac{price}{inc} + \beta_3 \frac{educ}{inc} + \beta_4 \frac{female}{inc} + \frac{u}{inc}$$

8.7 For the model:

$$y_{ie} = \beta_0 + \beta_1 x_{i,e,1} + \beta_2 x_{i,e,2} + ... + \beta_k x_{i,e,k} + u_i, \text{ where } u_i = f_i + v_{i,e}$$

i $Var(u_i) = V(f_i + v_{i,e}) = V(f_i) + V(v_{i,e}) + 2Cov(f_i, v_{i,e})$
Since $Cov(f_i, v_{i,e}) = 0$, then:
$Var(u_i) = V(f_i) + V(v_{i,e}) = \sigma_f^2 + \sigma_v^2$

ii $Cov(u_{i,e}, u_{i,g}) = Cov(f_i + v_{i,e}, f_i + v_{i,g})$
$Cov(u_{i,e}, u_{i,g}) = V(f_i) + Cov(f_i, v_{i,g}) + Cov(f_i, v_{i,e}) + Cov(v_{i,e}, v_{i,g})$
Since $\forall i \neq j,\ Cov(f_i, v_{i,j}) = 0$ and $Cov(v_{i,e}, v_{i,j})$, then:
$Cov(u_{i,e}, u_{i,g}) = V(f_i) = \sigma_f^2$

10.1 i This statement is most often not true. For the same individual (or group) observed repeatedly across time, the unobserved determinants of its outcomes are most likely correlated from one period to the next.

ii As indicated by Theorem 10.1, this statement is true.

10.5 There are several acceptable answers here:

If we have enough observations per quarter, one straightforward approach would be to create fixed effects for every year-quarter combination in the dataset, so that we have:

$$starts = \tau_1 yr1q1 + ...\tau_4 yr1q4 + ...\tau_{4n} yrNq4 + \beta_1 rate_{i,t} + \beta_2 inc_{it} + u_{it}$$

*Per.Hjertstrand@ifn.se

†Andrew.Proctor@phdstudent.hhs.se

If instead, we are more limited in our data, we may want to specify simply a year trend and quarter fixed-effects:

$$starts = \alpha + \tau T + \delta_1 q1 + \delta_2 q2 + \delta_3 q3 + \delta_3 q3 + \delta_4 q4 \beta_1 rate_{i,t} + \beta_2 inc_{it} + u_{it}$$

13.6 Let 1990 be a fixed effect for 1990 and FL be an indicator for Florida. Then we have:

i

$$arrest = \alpha + \tau 1990 + \gamma FL + \beta(FL \times 1990) + u$$

ii As is always the case with empirical problems, there are many omitted factors you might want to address. Specifically though, in a fixed effect model we are concerned with factors that have a differential effect on the two states across the periods 1985-1990.

Some regional controls might be in order. Certainly, it is fair to wonder whether factors like the degree of police presence are either constant in the states over the years, are trend in common.

Additionally, one might easily imagine that there were differential trends in alcohol use in Florida and Georgia that motivated Florida to address the problem. Ideally, if we had both individual data about average alcohol use and could make the (potential strong) assumption that this was uncorrelated with the open container law, this would be a powerful control

Another feature that might vary over time and between states might be demographics - for example, of individuals aged 18-30 was increasing in Florida (and this age group is more likely to get caught with an open container), then this again would be a violation of the common trends. If we had individual data, we could control directly for these possible confounders.

iii Ultimately this changes very little about the estimation. We would still proceed using difference-in-difference estimation, only now our variables would be county-level measures (such as arrest rates and county means of demographic indicator).

Stock and Watson

a There are many such variables. Some include cognitive abilities (i.e. intelligence) and noncognitive abilities (e.g. social skills), motivation and effort, responsibility, etc.

b Time specific variables most notably suggest macroeconomic conditions, such as growth rate of the economy and inflation.

c By including person-specific and time-specific indicator variables.

d Unless we have variation in personal gender among the 10,000 workers during the 2000-2016 period, then we cannot as it would be capture by the person-specific effects.

e Again, there are many unobserved characteristics you could imagine but you should once again focus on those that vary for the same individual over the panel period. Some potential confounders include vocational training, health status and number of children (determinants of labor force of attachment) and employment industry/sector. Importantly, however, while each of these factor may be confounders in the fixed effects regression approach, they are also problematic as controls, as each may be endogenous to education and earnings.

One variable you might be very much tempted to control for is occupation/position, but this would be a textbook example of a 'bad control,' as much of the effect of education on earnings likely operates through the channel of yielding higher-paid positions.

f The unobservable determinants of earning are likely to have unequal variance between, for example, high wage and low-wage (or unattached) workers, contributing to heteroskedasticity.

Moreover, the unobserved determinants of earnings are likely to be persistent for the same individuals across different years, hence the error terms for individuals are likely to be correlated (autocorrelatation).

g Although you very well may be concerned about dynamic effects, lagged dependent variables are not acceptable in fixed effect. Strict exogeneity requires $Cov(u_{it}, x_{is} = 0 \; \forall s$ in $1...T$, which is violated.

To see this, take for instance, $s = t + 1$ and let $x_{i,s} = y_{i,s-1}$ (the lag dependent variable).
Then: $Cov(u_{it}, x_{i,s}) = Cov(u_{it}, y_{i,s-1}) = Cov(u_{it}, y_{i,t})$.
And since $y_{it} = \hat{y}_{it} + u_{it}$,
Then $Cov(u_{it}, y_{i,t}) = Cov(u_{it}, \hat{y}_{it} + u_{it}) = Cov(u_{it}, \hat{y}_{it}) + Var(u_{it}) \neq 0$