REVISION NOTES

* TERM 1, TOPIC 9: RHS ENDOGENOUS VARIABLES *

This topic includes:

CONTROLLING FOR ENDOGENEITY

- REGRESSION DISCONTINUITY
- RANDOMISED CONTROLLED TRIALS
- IV
  - ASSUMPTION OF INSTRUMENT EXOGENEITY
  - ASSUMPTION OF INSTRUMENT RELEVANCE

DETECTING ENDOGENEITY: HAUSMAN-WU TEST

Explanatory variables are *not* given so everything we have done so far with regressions is wrong. It was far too simple to say that years of schooling affected wages because we *choose* the years of schooling we want, we choose the type of job we want, we *choose* how long we work, etc. Thus, we need something more realistic.

Consider the model of exam performance y_{1i}:

$$\underset{y_{1i}}{perf_i} = \beta_0 + \beta_1 \underbrace{lecturesatt_i}_{\substack{y_{2i} \\ \text{(choice/endogenous)}}} + \beta_2 \underbrace{female_i}_{\substack{x_{1i} \\ \text{(exogenous)}}} + \beta_3 \underbrace{alevel_i}_{\substack{x_{2i} \\ \text{(exogenous)}}} + \varepsilon_{1i}$$

- If $\beta_1 = 0$, then essentially there is no effect of lecture attendance on exam performance, so you may as well not show up at all
- Instead, if $\beta_1 =$ very positive, then we are saying there is much value added from lectures
- ε_{1i} includes all the indeterminable data for which we have no dataset but still affects performance, e.g. $\varepsilon_{1i} = (\underset{+}{effort_i}, \underset{-}{poorgraspoflang_i}, \underset{+}{ability_i})$ – more effort and ability will lead to higher grades whilst the more disassociated you are with England the worse grades you will get

Consider the model of lectures attended y_{2i}:

$$\underbrace{lecturesatt_i}_{\substack{y_{2i} \\ \text{(choice/endogenous)}}} = \delta_0 + \delta_1 \underbrace{extracurr_i}_{\text{endogenous}} + \delta_2 \underset{\text{endogenous}}{dist_i} + \delta_3 \underset{\text{endogenous}}{9am_i} + \delta_4 \underset{\text{exogenous}}{wed_i} + \delta_5 \underbrace{health_i}_{\text{exogenous}} + \delta_6 \underset{\text{exogenous}}{fem_i} + \varepsilon_{2i}$$