MaC-metrics

Josh Angrist

MIT 14.661 (FALL 2017)

A Life's Identification Challenges

- The iconic empirical life-cycle labor supply function looks like this:

$$\ln h_{it} = \mu_i + \quad (\rho \quad r)t + \quad \ln w_{it} + u_{it} \tag{1}$$

  where $\quad = \frac{1}{2 \quad 1}$ and u_{it} is a "tacked-on" error

- We're after , the ISE, but it's not easily captured

  - Estimation of (1) isn't shovel-ready: the control variable μ_i isn't found in the CPS. A function of the marginal utility of wealth, this variable is negatively correlated with wages, w_{it}
  - We have limited data on hourly wages; instead, we work with average hourly earnings, $AHE_{it} \equiv \frac{y_{it}}{h_{it}}$. So we're naively regressing *hours worked* on $(hours\, worked)^{\quad 1}$; The results might not be pretty; rather, they're pretty negative!

B Problems and Solutions

- Analysis of covariance (deviations from means) or differencing kills the unobserved fixed effect (whew!)

- These transformations also aggravate the bias from our poorly measured wage variable. The bias here is worse than classical attenuation bias: the fact that mismeasured hours appears on both sides of the equation of interest induces a powerful negative term known as "division bias"

- We might instead try grouping strategies, as in Angrist (1990, 1991). This approach potentially kills the measurement error as well as the fixed effect. I like that a lot!

C Division Bias Details

Suppose the labor supply equation of our heart's desire is

$$\ln h^*_{it} = \alpha + \quad \ln w^*_{it} + u_{it} \tag{2}$$

For the purposes of this discussion, we'll start by assuming we'd be happy to estimate (2) by OLS.

The empirical supply function uses AHE with well-measured hours

$$w^*_{it} = \frac{y_{it}}{h^*_{it}},$$

where y_{it} is annual earnings. This is the hourly wage for those who are paid hourly, and its a notional time price for others. Either way, we assume this correctly-measured AHE is what consumers use to make work decisions.

In practice, however, hours are poorly measured:

$$h_{it} = h^*_{it} \quad {}_{it},$$

where ${}_{it}$ is proportional classical measurement error. Then

$$\ln h_{it} = \ln h^*_{it} + \eta_{it}, \tag{3}$$

where $\eta_{it} = \ln \quad {}_{it}$. This implies that

$$\ln w_{it} = \ln y_{it} \quad \ln h_{it} = \ln y_{it} \quad \ln h^*_{it} \quad \eta_{it} = \ln w^*_{it} \quad \eta_{it} \tag{4}$$

Substituting on both sides of (2), we now have

$$\ln h_{it} = \alpha + \quad (\ln w_{it} + \eta_{it}) + u_{it} + \eta_{it} = \alpha + \quad \ln w_{it} + \{u_{it} + (1 + \quad)\eta_{it}\} \tag{5}$$

Without worrying about the fixed effect, the OVB in OLS estimates of (5) is

$$OVB = \frac{Cov(\ln w^*_{it} \quad \eta_{it}, (1 + \quad)\eta_{it})}{{}^2_{\ln w}} = \frac{(1 + \quad) \, {}^2_{\eta}}{{}^2_{\ln w}}$$

which is big-time bad, even compared to the usual m.e. attenuation bias. (Note that $\frac{{}^2_{\eta}}{{}^2_{\ln w}}$ is one minus the signal-to-noise ratio for log wages.)

Analysis of covariance aggravates division bias

To kill the fixed effect, you might difference or deviate from means. Suppose you have a two-period panel, so (2) with fixed effects becomes OLS on first diffs:

$$\ln h^*_{it} = \quad \ln w^*_{it} + \quad u_{it},$$

while the noisy wage becomes

$$\ln w_{it} = \quad \ln w^*_{it} \quad \quad \eta_{it}.$$

Assuming m.e. is serially uncorrelated, the variance of η_{it} is $2\ {}_{\eta}^{2}$.

Wages, by contrast, are highly persistent. Suppose, $w_{it}^{*} = w_{i}^{*}$. Then

$$\ln w_{it} = \quad \eta_{it}.$$

In other words, the change in wages is pure noise. Then we have

$$OVB = \frac{(1+\quad)2\ {}_{\eta}^{2}}{2\ {}_{\eta}^{2}} = \quad (1+\quad)$$

so differencing here makes matters substantially worse. Research on measurement error in hours and wages bears this out: measured wage changes are noisy indeed (see, e.g., Bound and Krueger, 1991).

See MaCurdy, Thomas E. "An Empirical Model of Labor Supply in a Life-Cycle Setting." *Journal of Political Economy* 89, no. 6 (1981): 1076-1078.

See Angrist, Joshua D. "Grouped-data estimation and testing in simple labor-supply models." *Journal of Econometrics* 47, no 2-3 (1991): 256-259.

See Card, David. "Intertemporal Labour Supply: An Assessment." In *Advances in Econometrics: Sixth World Congress, Vol II*. Cambridge University Press, 1996. pp. 59-60.