Econometrics II

Jan Toczynski

April 19, 2019

1. Answer to question 1

a) For a randomly chosen individual we have:

$$E(\ln w_i^U - \ln w_i^N | x_i^{'}) = x_i^{'}(\beta^U - \beta^N) + \underbrace{u_i^U - u_i^N}_{0} \tag{1}$$

b) For an individual from a union group we have:

$$\begin{aligned} &E(\ln w_i^U - \ln w_i^N | x_i^{'}, U_i^* > 0) = x_i^{'}(\beta^U - \beta^N) + E(u_i^U - u_i^N | U_i^* > 0) = \\ &= x_i^{'}(\beta^U - \beta^N) + E(u_i^U - u_i^N | \delta_0 + \delta_1(\ln w_i^U - \ln w_i^N) + \delta_2 x_i^{'} + \delta_3 z_i^{'} - v_i > 0) \\ &= x_i^{'}(\beta^U - \beta^N) + E(u_i^U - u_i^N | u_i^U - u_i^N - v_i > c) \end{aligned} \tag{2}$$

Now, we say:

$$u_i^U - u_i^N = a(u_i^U - u_i^N - v_i) + \epsilon_i \tag{3}$$

and a is:

$$a = \frac{Cov(u_i^U - u_i^N, u_i^U - u_i^N - v_i)}{Var(u_i^U - u_i^N)} \tag{4}$$

as the error terms are uncorrelated with each other, I drop the epsilon part when plugging the above to the first equation and I get:

$$\begin{aligned} &E(\ln w_i^U - \ln w_i^N | x_i^{'}, U_i^* > 0) = \\ &= x_i^{'}(\beta^U - \beta^N) + aE(u_i^U - u_i^N - v_i | u_i^U - u_i^N - v_i > c) = \\ &= x_i^{'}(\beta^U - \beta^N) + a\sqrt{\sigma_U^2 - \sigma_N^2 - \sigma_V^2}\lambda(c) \end{aligned} \tag{5}$$

as the errors are normally distributed and uncorrelated and thus their sum is a random normal variable.

b) No, it's not possible as the treatment status (being unionized) is correlated with the error terms so the estimates would be biased.

d) In the first stage we run a probit of:

$$\begin{aligned} U_i^* &= \delta_0 + \delta_1(\ln w_i^U - \ln w_i^N) + \delta_2 x_i^{'} + \delta_3 z_i^{'} - v_i = \\ &= \delta_0^* + \delta_1^* x_i^{'} + \delta_3^* z_i^{'} - \epsilon_i^* \end{aligned} \tag{6}$$

where $\epsilon_i^* = u_i^U - u_i^N - v_i$. This will give us consistent estimates of $\delta_0^*, \delta_1^*, \delta_2^*$ In the second stage, we use the above estimates in estimating individual wage equations. For example, for the first one that will be:

$$\ln w_i^U = \beta^U x_i' + \sigma_{1\epsilon^*}(-\frac{f(\psi_i)}{\Phi_i(\psi_i)}) + \nu_i \tag{7}$$

where $\psi_i = \hat{\delta_0^*} + \hat{\delta_1^*} x_i' + \hat{\delta_3^*} z_i'$, and the sigma-term, as in the paper, comes from normalizing the errors. According to the paper, this will give us consistent estimates of betas.

e)

f) The relative rate of return to education is higher in the non-unionized sector. The rest of the coefficients suggests that 1) being black or a woman has negative impact on wages 2) the effect of market experience is positive, but non-linear (decreasing marginal gain) 3) health problems have negative impact on wages.

g) The positive coefficient on the inverse Mills ratio coefficient can be interpreted as a positive selection - if we did not correct for selection, the coefficient would be biased upwards. The interpretation for the negative selectivity variable is analogous.

h) From table 6 we learn that in the structural equation, the coefficient of the wage differential is positive and statistically significant. We can interpret it as an indication that wage differential has strong impact on the decision to join a union.

2. Answer to question 2

a) See stata code attached.

b) Estimating the two probits jointly allows us to account for correlation between the two choice decisions.

c) We should estimate the following model:

$$P(\text{kidslt6} = 1|X) = \Phi(X\beta) \tag{8}$$

$$P(\text{inlf} = 1|X) = \Phi(X\beta + \alpha\text{kidslt6}) \tag{9}$$

In doing so, we assume that the error terms are jointly normal distributed with zero mean. The estimated ρ is around -0.7 (statistically significant). We can interpret this result as a presence of selection on unobservables.

d) The method relies on estimating the bivariate probit model with different constraints on the value of the correlation between error terms. One can treat $\rho = 0$ as a baseline case, in which there is no selection on unobservables and compare the coefficients of the model for different values of ρ. As an indication as to what values of ρ are plausible, one can try to bound them by making assumptions regarding how strong selection on unobservables is believed to be vis a vis selection on observables.

e) See code. Treating $\rho = 0$ as the baseline case, I get a coefficient of around -0.6 and highly significant. $\rho = 0.3$ already eliminates the significance, making the coefficient statistically indistinguishable from 0.

f) The bound we are after is $\rho = \frac{cov(X'\beta, X'\gamma)}{var(X'\gamma)}$ from the paper (here, I'm a bit confused, should β and γ come from the estimated model assuming $\rho = 0$ or rather from the estimation with no constraints?). For the latter I find the bound to be around -0.26, which is close to the value from e).

g) The estimated coefficient is around -0.22.