Econometrics 1

TD 6: linear regressions and causality

(Chapter 4, Section 1)

We consider a binary treatment variable D and the corresponding potential outcome variables $Y(0), Y(1)$, which are assumed to be such that

$$(Y(0), Y(1))' \sim \mathcal{N}(\mu, \Sigma)$$

with $\mu = (\mu_0, \mu_1)'$ the 2×1 vector of expectations and $\Sigma = (\Sigma_{jk})_{1 \leq j,k \leq 2}$ the 2×2 variance-covariance matrix. Hence, $(Y(0), Y(1))'$ is a Gaussian vector.

We consider two models for the treatment D:

- Model 1: $D \,|\, Y(0), Y(1) \sim \text{Bernoulli}(0.5)$;
- Model 2: $D = \mathbb{1}\{Y(1) - Y(0) > c\}$ for some non-stochastic constant $c \in \mathbb{R}$.

Hereafter, Δ, δ, δ^T, β_0, and B have the same meaning as in the course (see Chapter 4).

1. Interpret the two models and give an example where Model 1 holds. What is δ equal to in the two models?
2. Express δ^T and β_0 in Model 1.
3. In Model 2, what happens if $\Sigma_{11} = \Sigma_{12} = \Sigma_{22}$? Compute δ^T in this case.
4. Show, with almost no computation, that $\delta^T \geq \delta$ in Model 2.
5. Show that, for some (a, b) to be specified,

   $$Y(0) = a + b\Delta + \xi, \quad \text{with } E(\xi) = 0 \text{ and } \xi \perp\!\!\!\perp \Delta.$$

   Deduce the sign of the selection bias B in Model 2 as a function of Σ. Interpret the results.
6. Check through simulations (with the parameters $\mu = (0, 1)'$, $\Sigma =$ identity matrix and $c = 1.5$) the value of δ in the two models.
7. Estimate, through simulations again, the values of δ^T and β_0 in Model 2. If time permits, check that your simulations are correct by doing analytical computations.