The Roy Model: Basics

Applied Microeconomics

Hans-Martin von Gaudecker & Florian Zimmermann

The (Generalised) Roy Model

- Indirect utility from choosing s when Z takes on value z:

$$R(s, z, \omega)$$

- Assume additive separability:

$$R(s, z, \omega) = \mu(s, z) - \eta(s, \omega)$$

- Assume utility maximisation:

$$S = \arg\max_{s \in \mathcal{S}} R(s, z, \omega)$$

The Generalised Roy Model

- Only utility differences between choices matter
- Binary case:

$$S = \mathbb{I}[\tilde{\eta}(\omega) \leq \tilde{\mu}(z)]$$

with $\tilde{\eta}(\omega) \equiv \eta(1, \omega) - \eta(0, \omega)$ and $\tilde{\mu}(z) \equiv \mu(1, z) - \mu(0, z)$

A very useful representation

- Integrate over the distribution of $\tilde{\eta}(\omega)$ on both sides:

$$\begin{aligned} S &= \mathbb{I}[\tilde{\eta}(\omega) \leq \tilde{\mu}(z)] \\ &= \mathbb{I}[F_{\tilde{\eta}}\left(\tilde{\eta}(\omega)\right) \leq F_{\tilde{\eta}}\left(\tilde{\mu}(z)\right)] \\ &\equiv \mathbb{I}[U \leq P\left[S = 1 | Z = 1\right]] \end{aligned}$$

with $U \sim \mathrm{Uniform}(0,1)$

- Only a normalisation step
- Propensity to take treatment

Resulting equations for $E[Y(S, \omega)|U = u]$

- Conditional expectations of Y depending on treatment S and the unobservable component of utility being at its u'th quantile:

$$E[Y(S = 0, \omega)|U = u] = m_0(u)$$
$$E[Y(S = 1, \omega)|U = u] = m_1(u)$$

- "Marginal treatment response functions"

Non-compliance × 1 or 2

- One-sided non-compliance motivated early literature on selection models
  - Only observe wages of those who work
  - Only observe data for those who respond to survey
- Standard treatment-effect model has two-sided non-compliance
  - Some choose not get treated who are encouraged (``never-takers")
  - Some choose treatment despite not being encouraged (``always-takers")
  - In binary case:

$$0 < P[S = 1|Z = 0] < P[S = 1|Z = 1] < 1$$

A very useful result on LATE

Kline and Walters ("On Heckits, Late, And Numerical Equivalence", 2019), Theorem 1:

Under two-sided non-compliance, 2SLS estimation of LATE is equivalent to estimation using parametric assumptions on the MTR functions $m_0(u)$ and $m_1(u)$ of the form $\alpha_s + \gamma_s \cdot (J(u) - E[J(U)])$, with $J(u)$ being strictly increasing.