The Generalized Roy Model - TA Session

Nadav Kunievsky

January 12, 2021

What type of questions we tackle in economics? (Heckman, 2010)

Three broad classes of problems we consider in economics

P1: Evaluating the impact of historical interventions on outcomes including their impact in terms of the well-being of the treated and society at large

P2: Forecasting the impacts (constructing counterfactual states) of interventions implemented in one environment in other environments, including their impacts in terms of well-being

P3: Forecasting the impacts of interventions (constructing counterfactual states associated with interventions) that were never historically experienced to various environments, including their impacts in terms of well-being

Advantages of the Roy Model as an empirical framework

- Allows to address P1-P2
- Model explicitly what are the policy invariant parameters
- Defines the choice mechanism and treatment assignment rule
- Marschak's Maxim/Occam's Razor - allows to reduce the needed assumptions
- Brings forth the a discussion on subjective Vs. Objective utilities /Ex-ante Vs. Ex-post

The Basic Roy Model

- Consider the following model
  - Workers have two types of skills S_1^i, S_2^i.
  - Skill premiums are given by π_1 and π_2.
  - Agents choose to work where there earnings are the highest

    $$\pi_1 S_1^i \geq \pi_2 S_2^i$$

    (Assuming no ties).

The Basic Roy Model

- Given some distribution of skills, we can derive the following
  - The share of workers in sector 1 is given by

$$P_1 = \Pr(\pi_1 S_1 > \pi_2 S_2) = \int_0^{\infty} \int_0^{\pi_1 s_1/\pi_2} f(s_1, s_2)\, ds_2 ds_1$$

  - The density of workers with s_1 is different from the share of skilled workers in the population (Selection)

$$f_1^p(s_1) = \int_0^{\infty} f(s_1, s_2)\, ds_2$$

$$g(s_1 \mid \pi_1 s_1 > \pi_2 s_2) = \frac{1}{P_1} \int_0^{\pi_1 s_1/\pi_2} f(s_1, s_2)\, ds_2$$

  - The distribution of wages in sector 1 (Heterogeneity in Outcomes)

$$g(w_1) = \frac{1}{P_1} \int_0^{\frac{2_1}{\pi_2}} f(w_1, s_2) ds_2$$

The Basic Roy Model - Adding Normality

- Assume log skills are Normally distributed

$$\ln S_j = \mu_j + U_j$$

$$\begin{pmatrix} U_1 \\ U_2 \end{pmatrix} \sim N\left(\begin{matrix} 0 \\ 0 \end{matrix}, \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} \right)$$

$$\implies \ln W_j = \ln \pi_j + \mu_j + U_j, j = 1, 2$$

The Basic Roy Model - Adding Normality

- Given the normality assumption we can solve for the explicitly for the expressions we want

  $$E(lnW_1|lnW_1 \geq lnW_2) =$$
  $$\log \pi_1 + \mu_1 + E(U_1|U_1 - U_2 \geq (log\pi_2 - \mu_2) - (log\pi_1 + \mu_1))$$

  Notice that both $U_1 - U_2$ and U_1, are jointly distributed normally with

  $$\begin{pmatrix} U_1 \\ U_1 - U_2 \end{pmatrix} \sim N\left(\begin{matrix} 0 \\ 0 \end{matrix}, \begin{bmatrix} \sigma_{11} & \sigma_{11} - \sigma_{12} \\ \sigma_{11} - \sigma_{12} & \sigma_{21} + \sigma_{22} - 2\sigma_{12} \end{bmatrix} \right)$$

- Where we used the normality assumption and the fact that

  $$\mathrm{Var}(U_1 - U_2) = \sigma_{11} + \sigma_{22} - 2\sigma_{12}$$
  $$\mathrm{Cov}(U_1, U_1 - U_2) = \sigma_{11} - \sigma_{12}$$

The Basic Roy Model - Adding Normality

- We can now use the properties of the normal to derive an explicit expression of the average wage in sector 1
- let $C = (log\pi_2 - \mu_2) - (log\pi_1 + \mu_1)$, then

$$E(U_1|U_1 - U_2 \geq C) = \frac{\mathrm{Cov}(U_1, U_1 - U_2)}{\mathrm{Var}(U_1 - U_2)} E[(U_1 - U_2)|U_1 - U_2 \geq C$$

$$= \frac{\sigma_{11} - \sigma_{12}}{\sqrt{\sigma_{11} + \sigma_{22} - 2\sigma_{12}}} \underbrace{\frac{\phi(C)}{1 - \Phi(C)}}_{\text{Inverse Mills Ratio}}$$

- where we used the population regression

$$U_1 = \beta(U_1 - U_2) + \epsilon = \frac{\mathrm{Cov}(U_1, U_1 - U_2)}{\mathrm{Var}(U_1 - U_2)}(U_1 - U_2) + \epsilon$$

and the expectation of the truncated normal distribution with mean μ

$$\mathrm{E}(X \mid a < X < b) = \mu + \sigma \frac{\phi(\alpha) - \phi(\beta)}{\Phi(\beta) - \Phi(\alpha)}$$

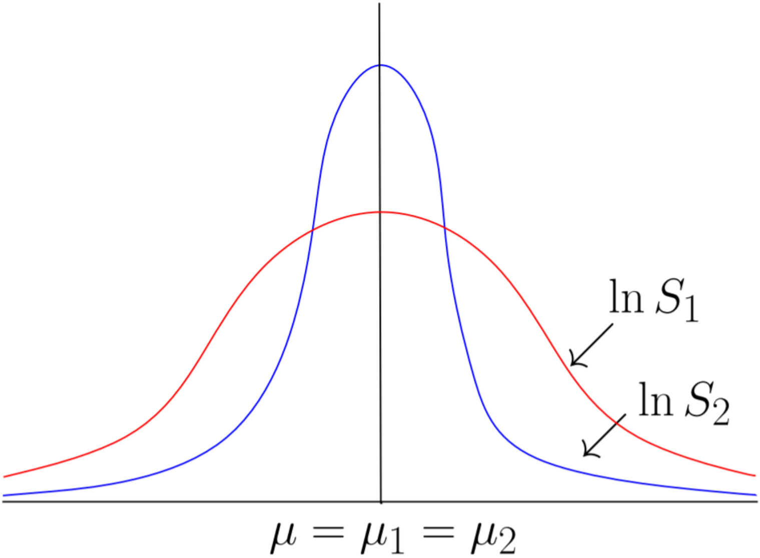
The Basic Roy Model - Adding Normality

- Therefore the average wage in the sector is given by

$$E(w_1|w_1 \geq w_2) =$$
$$\log \pi_1 + \mu_1 + \frac{\sigma_{11} - \sigma_{12}}{\sqrt{\sigma_{11} + \sigma_{22} - 2\sigma_{12}}} \frac{\phi(C)}{1 - \Phi(C)}$$

- Selection component of the average wage in sector 1 is driven by the correlation between U_1 and $U_1 - U_2$, and the difference in prices (C)
- Notice that we could have negative selection, where people with higher skill are less likely to go into the sector. This happens when $\sigma_{11} - \sigma_{21} < 0$.
- Selection can play a significant role when we compare averages
- The selection equation $U_1 - U_2 \geq C$, conveys information on what the agent acts upon (Subjective Vs. Objective utilities, Ex-ante Vs. Ex-post benefits).

Example - $\pi_1 = \pi_2, \mu_1 = \mu_2$



Densities of $\ln S_1$ and $\ln S_2$

THE UNIVERSITY

The Generalized Roy Model

- We can generalize the basic Roy model, by considering general outcomes, and adding additional cost shifters

$$
\begin{aligned}
Y_1 &= \mu_1(X) + U_1 \\
Y_0 &= \mu_0(X) + U_0 \\
C &= \mu_C(Z) + U_C \\
I &= Y_1 - Y_0 - C \implies \\
I &= \underbrace{\mu_1(X) - \mu_0(X) - \mu_C(Z)}_{\mu_D(Z)} + \underbrace{U_1 - U_0 - U_C}_{-V}
\end{aligned}
$$

$$
\begin{aligned}
&(U_0, U_1, U_C)\,(X, Z) \\
&E\,(U_0, U_1, U_C) = (0, 0, 0) \\
&V \perp\!\!\!\perp (X, Z)
\end{aligned}
$$

The Generalized Roy Model

- The econometrician observes

$$Y = DY_1 + (1 - D)Y_0$$
$$D = 1(I \geq 0) = 1\left(\mu_D(Z) \geq V\right)$$

- Notice that the selection equation, the propensity score, is a function of Z (conditional on X). This hints for non-parametric estimation. In the unique case in which V is normally distributed we have

$$\Pr(D = 1 \mid Z = z) = \Phi\left(\frac{\mu_D(z)}{\sigma_V}\right)$$

- The observed outcome, for the treated, is given by

$$E(Y \mid D = 1, X = x, Z = z) = \mu_1(X) + \underbrace{E\left(U_1 \mid \mu_D(z) \geq V\right)}_{K_1(P(z))}$$

The Generalized Roy Model

- Under normality, as we've seen, we have

$$E(Y \mid D = 1, X = x, Z = z) = \mu_1(x) + \left(\frac{\text{Cov}\left(U_1, \frac{V}{\sigma_V}\right)}{\text{Var}\left(\frac{V}{\sigma_V}\right)} \right) \tilde{\lambda} \left(\frac{\mu_D(z}{\sigma_V} \right.$$

$$E(Y \mid D = 0, X = x, Z = z) = \mu_0(x) + \left(\frac{\text{Cov}\left(U_0, \frac{V}{\sigma_V}\right)}{\text{Var}\left(\frac{V}{\sigma_V}\right)} \right) \lambda \left(\frac{\mu_D(z}{\sigma_V} \right.$$

- If we are willing to assume normality we can identify the model parameters (and therefore answer causal questions), by MLE or a two-step method. If we are not willing to assume normality, then, under some assumptions, we can identify
  - $P(z)$ can be estimated non-paramatrically
  - Given $P(z)$, we can identify $\mu_1(X)$ and $\mu_0(X)$ from the conditional expectations
  - To identify the U's distribution we can use identification at infinity, see "Notes on Identification of the Roy Model and the Generalized Roy Model"

The Generalized Roy Model For Policy Evaluations

- The Roy Model framework, allows us to identify parameters of interest to evaluate policies
- The Average Treatment Effect is

$$E[Y_1 - Y_0|X] = \mu_1(x) - \mu_0(x)$$

The Generalized Roy Model For Policy Evaluations

- We can derive the treatment on the treated, $E(Y_1 - Y_0|D = 1, X, Z)$, using

$$
\begin{aligned}
E\left(Y_1 \mid D = 1, X, Z\right) = \mu_1(x) + K_1(P(z)) \\
E\left(Y_0 \mid D = 0, X, Z\right) = \mu_0(x) + \tilde{K}_0(P(z))
\end{aligned}
$$

where

$$
\begin{aligned}
K_1(P(z)) = E\left(U_1 \mid \frac{\mu_D(z)}{\sigma_V} > \frac{V}{\sigma_V}\right) \\
\tilde{K}_0(P(z)) = E\left(U_0 \mid \frac{\mu_D(z)}{\sigma_V} \leq \frac{V}{\sigma_V}\right)
\end{aligned}
$$

And using the fact that $E(U_1) = E(U_0) = 0$, which gives us

$$K_1(P(z))P(z) + \tilde{K}_1(P(z))(1 - P(z)) = 0$$

$$(1 - P(z))\tilde{K}_0(P(z)) + P(z)K_0(P(z)) = 0$$

The Generalized Roy Model For Policy Evaluations

- Combining these expressions we get

$$E(Y_1 - Y_0|D = 1, x, z) = \mu_1(x) - \mu_0(x) + K_1(P(z)) - K_0(P(z))$$

- Similarly, we can construct the ATU

$$E(Y_1 - Y_0|D = 0, x, z) = \mu_1(x) - \mu_0(x) + \tilde{K}_1(P(z)) - \tilde{K}_0(P(z))$$

MTE

- We can also identify the treatment effect on the people at the margins

$$E\left(Y_1 - Y_0 \mid I = 0, X = x, Z = z\right) \\ = \mu_1(x) - \mu_0(x) + E\left(U_1 - U_0 \mid \tfrac{\mu_D(z)}{\sigma_V} = \tfrac{V}{\sigma_V}, X = x, Z = z\right)$$

- Where in the normal case we get

$$MTE(v) = \mu_1(x) - \mu_0(x) + \text{Cov}\left(U_1 - U_0, \frac{V}{\sigma_V}\right) v$$

- and at the margin

$$MTE(v) = \mu_1(x) - \mu_0(x) + \text{Cov}\left(U_1 - U_0, \frac{V}{\sigma_V}\right) \frac{\mu_D(z)}{\sigma_V}$$

MTE

- As shown in Heckman, 2010 (and other papers by Heckman and Vytlacil), the MTE can be used as a building block to all other causal parameters such as ATE, ATT, LATE, policy relevant treatment effect.
- It relates to the Marginal Revolution in econ, where it allows us to as, what the marginal benefit from a policy