Program Evaluation(b): Parametric Selection

Chris Conlon

March 28, 2021

Applied Econometrics

Example: Borjas (1987)

- Consider two countries $(0/1)$ (source and host).

$$\ln w_0 = \alpha_0 + u_0 \quad \text{with } u_0 \sim N(0, \sigma_0^2) \text{ source country}$$
$$\ln w_1 = \alpha_1 + u_1 \quad \text{with } u_1 \sim N(0, \sigma_1^2) \text{ host country}$$

- Now we allow for migration cost of C which he writes in hours: $\pi = \frac{C}{w_0}$.
- Assume workers know everything; you only see u_0 OR u_1 depending on country.
- Correlation in earnings is $\rho = \frac{\sigma_{01}}{\sigma_0 \sigma_1}$.

Example: Borjas (1987)

- Workers will migrate if:

$$(\alpha_1 - \alpha_0 - \pi) + (u_1 - u_0) > 0$$

- Who migrates? Probability of migration. Define $\nu = u_1 - u_0$.

$$\begin{aligned} P = \Pr\left[\nu > (\alpha_0 - \alpha_1 + \pi)\right] &= \Pr\left[\frac{\nu}{\sigma_\nu} > \frac{(\alpha_0 - \alpha_1 + \pi)}{\sigma_\nu}\right] \\ &= 1 - \Phi\left(\frac{(\alpha_0 - \alpha_1 + \pi)}{\sigma_\nu}\right) \equiv 1 - \Phi(z) \end{aligned}$$

- Higher $z \rightarrow$ less migration.

Example: Borjas (1987): How does selection work?

Construct counterfactual wages for workers in source country for those who immigrate:

- For now ignore mean differences $\alpha_0 = \alpha_1 = \alpha$.

$$\begin{aligned}\mathbb{E}\left(w_0 | \text{ Immigrate }\right) &= \alpha + \mathbb{E}\left(u_0 | \frac{\nu}{\sigma_\nu} > z\right) \\ &= \alpha + \sigma_0 \cdot \mathbb{E}\left(\frac{u_0}{\sigma_0} | \frac{\nu}{\sigma_\nu} > z\right)\end{aligned}$$

- Wages depend on:
  1. Mean earnings in the source country
  2. Both error terms (u_0, u_1) through ν
  3. Implicitly, it also depends on the correlation between the error terms.

Example: Borjas (1987): How does selection work?

- If everything is normal, we just run univariate regression $\mathbb{E}\left(u_0|\nu\right) = \frac{\sigma_{0\nu}}{\sigma_\nu^2}\nu$:

$$\mathbb{E}\left(\frac{u_0}{\sigma_0}|\frac{\nu}{\sigma_\nu}\right) = \frac{1}{\sigma_0}\cdot\frac{\sigma_{0\nu}}{\sigma_\nu^2}\cdot\frac{\sigma_\nu^2}{\sigma_\nu^2}\cdot\nu = \frac{\sigma_{0\nu}}{\sigma_0\sigma_\nu}\frac{\nu}{\sigma_\nu} = \rho_{0\nu}\frac{\nu}{\sigma_\nu}$$

$$\begin{aligned}\mathbb{E}\left(w_0|\text{ Immigrate }\right) &= \alpha_0 + \sigma_0\cdot\mathbb{E}\left(\frac{u_0}{\sigma_0}|\frac{\nu}{\sigma_\nu} > z\right)\\ &= \alpha_0 + \rho_{0\nu}\cdot\sigma_0\cdot\mathbb{E}\left(\frac{\nu}{\sigma_\nu}|\frac{\nu}{\sigma_\nu} > z\right)\\ &= \alpha_0 + \rho_{0\nu}\cdot\sigma_0\left(\frac{\phi(z)}{1-\Phi(z)}\right)\end{aligned}$$

- This hazard rate of the standard normal has a special name Inverse Mills Ratio $\mathbb{E}[x|x > z]$.

Example: Borjas (1987): How does selection work?

- A similar expression for those who do immigrate:

$$\begin{aligned}\mathbb{E}\left(w_1 \middle| \text{ Immigrate }\right) &= \alpha_1 + \mathbb{E}\left(u_1 \middle| \frac{\nu}{\sigma_\nu} > z\right) \\ &= \alpha_1 + \rho_{1\nu}\sigma_1 \left(\frac{\phi(z)}{\Phi(-z)}\right)\end{aligned}$$

- We can re-write both expressions in terms of the Inverse Mills Ratio

Inverse Mills Ratio

$$
\begin{aligned}
\mathbb{E}\left(w_0 \middle| \text{ Immigrate }\right) &= \alpha_0 + \rho_{0\nu}\sigma_0 \left(\frac{\phi(z)}{1-\Phi(z)}\right) \\
&= \alpha_0 + \frac{\sigma_0\sigma_1}{\sigma_\nu}\left(\rho - \frac{\sigma_0}{\sigma_1}\right)\left(\frac{\phi(z)}{1-\Phi(z)}\right) \\
\mathbb{E}\left(w_1 \middle| \text{ Immigrate }\right) &= \alpha_1 + \rho_{1\nu}\sigma_1 \left(\frac{\phi(z)}{1-\Phi(z)}\right) \\
&= \alpha_1 + \frac{\sigma_0\sigma_1}{\sigma_\nu}\left(\frac{\sigma_1}{\sigma_0} - \rho\right)\left(\frac{\phi(z)}{1-\Phi(z)}\right)
\end{aligned}
$$

Where $\rho = \sigma_{01}/\sigma_0\sigma_1$.

Positive Hierarchical Sorting

Let $Q_0 = E\left(u_0 | I = 1\right), Q_1 = E\left(u_1 | I = 1\right)$ (expected skill of immigrants).

- Immigrants are positively selected and above average $(Q_0, Q_1) > 0$ and $\frac{\sigma_1}{\sigma_0} > 1$ and $\rho > \frac{\sigma_0}{\sigma_1}$
  - $\frac{\sigma_1}{\sigma_0} > 1$ returns to "skill" are higher in host country.
  - $\rho > \frac{\sigma_0}{\sigma_1}$ correlation between valued skills in both counties is high (similar skills valued in both countries).
- Best and brightest leave because returns to skill are too low in home country.

Negative Hierarchical Sorting

We swap the standard deviations:

- Immigrants are negatively selected and below average $(Q_0, Q_1) < 0$ and $\frac{\sigma_1}{\sigma_0} > 1$ and $\rho > \frac{\sigma_0}{\sigma_1}$
  - $\frac{\sigma_0}{\sigma_1} > 1$ returns to "skill" are lower in host country.
  - $\rho > \frac{\sigma_1}{\sigma_0}$ correlation between valued skills in both counties is high (similar skills valued in both countries).
- Compressed wage structure attracts the low skill types because it provides "insurance" or "subsidizes" low wage workers.

Refugee/Superman Sorting?

- Immigrants are below average at home and above average in host $(Q_0 < 0, Q_1 > 1)$ and $\frac{\sigma_1}{\sigma_0} > 1$:
  - $\rho < \min\left(\frac{\sigma_1}{\sigma_0}, \frac{\sigma_0}{\sigma_1}\right)$ being below average in source country makes you above average in host country.
- You are a nerdy intellectual in a country that values physical labor, or are otherwise discriminated against in the labor market.

The missing (fourth) case:

- Mathematically impossible $\rho > \max\left(\frac{\sigma_1}{\sigma_0}, \frac{\sigma_0}{\sigma_1}\right)$

Takeaway

What can we learn here?

- Heckman won a Nobel Prize for his work on selection...
- You need to know what an inverse Mills ratio is
- But today it is hard to get away with strong parametric assumptions (bivariate normal) on error terms.
- Doing MLE with a fully normal model is not a terrible place to start sometimes
    - Sometimes helpful to know how bad the selection problem might be.
- R package is `sampleSelection` and see `https://rpubs.com/hacamvan/316839` and `https://cran.r-project.org/web/packages/sampleSelection/vignettes/selection.pdf`.