

附录：内生参数识别与估计

根据前文假定区域内迁移不存在迁移成本且迁移成本是对称的，即 $\tau_{oo} = \tau_{dd} = 0$ 且 $\tau_{do} = \tau_{od}$ 。此外，为便于估计，只对部分无量纲的参数进行标准化，如劳动生产率 w 、受教育水平 q 和宜居条件 α ，即令 $w_1 = 1$ 、 $q_1 = 1$ 以及 $\alpha_1 = 1$ 。首先，我们对理论模型中劳动者平均工资水平 $\overline{inc}_{do}$ 方程、劳动者区位选择方程 π_{do} 等进行变换，用计量方法对内生参数 θ 、 ρ 、 w 、 q 、 τ 和 α 等进行估计。

在估计参数之前，本文需要对相关模型进行变换以便于参数识别。

首先，根据劳动者平均工资水平 $\overline{inc}_{do}$ 方程，对其进行对数变换，可以得到参数估计方程：

$$\ln(\overline{inc}_{do}) = \underbrace{\ln(\bar{\Gamma}) + \ln(w_d)}_{\text{工作地固定效应}} - (1/\theta) \ln(\pi_{do}) + \underbrace{\ln(q_o)}_{\text{出生地固定效应}} + \ln(\epsilon_{do}^w) \quad (1)$$

其次，根据劳动者区位选择方程 π_{do} ，对其进行对数变换，可以得到参数估计方程：

$$\ln(\pi_{do}) = \theta \ln(w_d) + \theta \ln(\alpha_d) + \theta \ln(1 - \tau_{do}) - \ln\left(\sum_{j=1}^N \tilde{w}_{jo}^\theta\right) + \theta \left[\ln(\epsilon_{do}^w) + \ln(\epsilon_{do}^\alpha)\right] \quad (2)$$

(1) 对 θ 和 q_o 的估计。方程 (1) 可以分离为目的地固定效应和出生地固定效应两部分，其中目的地固定效应由劳动生产率和技能参数组成，即 $DestinationFE_d = \ln \bar{\Gamma} + \ln w_d$ ，出生地固定效应为出生地的教育质量，即 $OrigionFE_o = \ln q_o$ 。通过平均工资 $\overline{inc}_{do}$ 对迁移人口比重 p_{do} 进行回归，控制出生地和目的地固定效应：

$$\ln \overline{inc}_{do} = DestinationFE_d + \beta \ln \pi_{do} + OrigionFE_o + \mu_{do}^w \quad (3)$$

即可计算出内生参数 $\theta = -1/\beta$ ， $q_o = \exp(OrigionFE_o)$ 。

(2) 对 ρ 和 w_d 的估计

利用 Fréchet 分布的性质有：

$$\frac{\text{Var}[inc_{ido}]}{(\overline{inc}_{do})^2} = \frac{E[inc_{ido}^2] - (E[inc_{ido}])^2}{(\overline{inc}_{do})^2} = \frac{\Gamma\left(1 - \frac{2}{\theta(1-\rho)}\right)}{\left[\Gamma\left(1 - \frac{1}{\theta(1-\rho)}\right)\right]^2} - 1 \quad (4)$$

上述方程中，劳动力工资的方差 $\text{Var}[inc_{ido}]$ 和均值 $\overline{inc}_{do}$ 以及技能异质性 θ 已知，求解该方程得到技能相关度参数 ρ 。结合技能异质性参数 θ 得到 $\bar{\Gamma}$ 的值，代入目的地固定效应，求出劳动生产率水平为 $w_d = \exp(DestinationFE_d - \ln \bar{\Gamma})$ 。

(3) 对 τ_{do} 的估计。对方程 (21) 进行变形得到迁移成本 τ_{do} 的估计方程：

$$\tau_{do} = 1 - \exp\left(\frac{\ln \pi_{do} - \ln \pi_{oo} + \ln \pi_{od} - \ln \pi_{dd}}{2\theta}\right) \quad (5)$$

(4) 对宜居程度 α_d 的估计。对方程 (2) 进行差分，可以消除 $\ln\left(\sum_{e=1}^N \Omega_{eo}^\theta\right)$ ，得到相对宜居程度 α_d ：

$$\ln \alpha_d = \frac{\sum_{o=1}^N \ln \pi_{do} - \sum_{o=1}^N \ln \pi_{1o}}{N\theta} - \frac{\sum_{o=1}^N \ln(1-\tau_{do}) - \sum_{o=1}^N \ln(1-\tau_{1o})}{N} - \ln w_d \quad (6)$$

其中 N 表示区域的数量，将迁移劳动比重 π_{do} 、迁移成本 τ_{do} 和技能异质性参数 θ 代入方程（6）就可以得到宜居程度 α_d 的估计值。