

Métodos de Evaluación de Impacto de Políticas Públicas

Trabajo Práctico

2014

El trabajo deberá ser entregado el día 27 de Agosto a Paula Salinardi antes de las 2:00 PM. Para realizar el trabajo debe utilizarse STATA. También deberán enviarse los archivos *.do* utilizados para llegar a los resultados a badaracconicolos@gmail.com. El objetivo de este trabajo es comparar la *performance* de distintos indicadores para estimar el impacto de un programa de entrenamiento para individuos con bajo capital humano. Este trabajo se basa en la simulación desarrollada en la sección 8.3 de Heckman, Lalonde y Smith (1999), para ello se generarán muestras de distribuciones con los parámetros que se definen a continuación.

Sea $D_i = 1$ si el individuo i participa del programa de entrenamiento y $D_i = 0$ en el caso contrario. Los ingresos del individuo i en el período t están determinados por:

$$Y_{it} = \beta + \alpha_i D_i + \theta_i + U_{it}$$

donde β es una constante, α_i es el efecto individual de participar en el programa, θ_i son efectos fijos por individuo y U_{it} un proceso autoregresivo de orden 1 definido como $U_{it} = \rho U_{it-1} + \varepsilon_{it}$, siendo los ε_{it} independientes e idénticamente distribuidos, con $E(\varepsilon_{it}) = 0$. Se asume también que $E(\theta_i) = 0$. Además, ε_{it} , α_i y θ_i son independientes. En términos de resultados potenciales, el modelo puede escribirse como:

$$Y_{it} = D_i Y_{1it} + (1 - D_i) Y_{0it} \quad Y_{1it} - Y_{0it} = \alpha_i$$

Los individuos deciden participar voluntariamente en el programa de entrenamiento. Los parámetros de interés de nuestro modelo son $E(\alpha_i)$ y $E(\alpha_i | D_i = 1)$. Existen T períodos de datos, donde k corresponde al período del programa. Asumiendo perfecta certidumbre, cada participante elige participar en el programa si el valor presente de los beneficios descontados supera los costos del entrenamiento. En términos de notación:

$$\begin{aligned} D_i &= 1 & \text{si } \alpha_i/r - Y_{ik} - c_i > 0 \text{ y } t > k \\ D_i &= 0 & \text{en cualquier otro caso,} \end{aligned}$$

donde r es la tasa de descuento y c_i los costos de participar en el programa, Y_{ik} es el costo de oportunidad en términos de ingresos (se asume que los participantes tienen ingreso cero en el período del programa). La función de costos se define como $c_i = Z_i \Phi + V_i$, en donde Z_i es una característica observada que afecta el costo de participar y V_i un costo no observado que asumimos se distribuye con media cero y varianza σ_V^2 . Además Z_i y V_i son independientes de todas las demás variables.

En lo que sigue, deberá simular 100 muestras de 1000 “individuos” cada una y para cada individuo generar 10 años de datos de ingresos: 5 años ($k-5$ a $k-1$) pre programa, un año $k = 6$ correspondiente en el cual se implementa el programa y los años $k+1$ a $k+4$ post programa.

En el año inicial ($k - 5$) se tiene que $U_{ik-5} = \varepsilon_{ik-5}$. Para la ecuación de ingresos asumir que $\beta = 1000$, el efecto tratamiento $\alpha_i \sim N(100, 300)^1$, $\theta_i \sim N(0, 300)$, y $\varepsilon_{it} \sim N(0, 280)$, mientras que $\rho = 0,78$. Para la función de costos suponer que $\Phi = 1$ y que las variables de costos presentan las siguientes distribuciones $V_i \sim N(0, 200)$ y $Z_i \sim N(\mu_Z, 300)$ donde μ_Z se fija de tal manera que en cada muestra el 10 % de los individuos participa en el programa. Suponga que la tasa de descuento es $r = 0,10$. Notar que $E(\alpha_i)$ y $E(\alpha_i|D_i = 1)$ se calculan sumando sobre las 100 observaciones.

Responda las siguientes preguntas:

1. ¿Cómo varía el costo de oportunidad de participar en el programa cuando aumenta θ_i ?
¿Qué sesgo tienen los estimadores de α_i mediante mínimos cuadrados?
2. ¿Bajo qué condiciones es Z_i un instrumento válido para D_i ?
3. ¿Existe “selección en el programa”? ¿Cuánto difieren $E(\alpha_i)$ y $E(\alpha_i|D_i = 1)$?
4. Calcule el estimador “cross section” con datos post programa para $k + 1$, $k + 2$, $k + 3$ y $k + 4$.
5. Calcule el estimador de diferencias en diferencias utilizando datos de los períodos $k - 1$ y $k + 3$, $k - 3$ y $k + 3$, $k - 5$ y $k + 3$. ¿Hay evidencia de “Ashenfelter dip”²? Comente los resultados. ¿Cuáles de los estimadores presentan mayores sesgos respecto del parámetro de interés?
6. Repita las estimaciones de impacto del programa para los períodos $k + 1$, $k + 2$, $k + 3$ y $k + 4$ utilizando Z_i como instrumento para la participación. ¿Cómo se interpreta el estimador con variables instrumentales?
7. Ahora suponga que $\alpha_i = \alpha = 100$ y repita las estimaciones de los puntos 4, 5 y 6. Comente los resultados en relación a los resultados de los puntos anteriores.
8. Construya un grupo de control realizando un *matching* con el vecino más cercano en términos de Y_{k-2} con reemplazo (antes de realizar las estimaciones) y repita las estimaciones de los puntos 4, 5 y 6, nuevamente suponiendo $\alpha_i = \alpha = 100$. Comente como cambian los resultados.

Referencias

HECKMAN, J., LALONDE, R. Y SMITH, J. (1999). *The economics and econometrics of active labor market programs*. Handbook of Labor Economics 3A.

ASHENFELTER, O. Y D. CARD (1985). “Using the longitudinal structure of earnings to estimate the effect of training programs.” *Review of Economics and Statistics* 67(3).

¹En el ejercicio 7 y 8 se asume que el efecto del programa es homogéneo entre los individuos, definido como $\alpha_i = \alpha = 100$.

²Ver Ashenfelter y Card (1985).