Trabajo Práctico- Evaluación de Impacto

Víctor Funes Leal

Pregunta 1

El parámetro θ_i representa al efecto específico de cada individuo, distribuido como normal con media 0 y varianza σ_θ^2, a su vez, éste parámetro determina el ingreso de todos los períodos y_{ik}. θ_i es también independiente de D_i, es decir: $cov(\theta_i, D_i) = 0$ por construcción, lo que implica que la decisión de participar en el programa dependerá de éste, pero solo a través de su relación con y_{ik} y por ello las personas con valores más altos de θ_i tenderán a no participar del programa porque su costo de oportunidad será más alto que para el resto, por lo que la tasa de participación será más baja que la que existiría si el θ_i fuese constante. La principal consecuencia de ésto es que el estimador de mínimos cuadrados de α_i estará sesgado hacia abajo.

El sesgo del estimador por mínimos cuadrados puede obtenerse de la siguiente forma: el modelo verdadero es el dado en el texto:

$$Y_{it} = \beta + \alpha D_i + \theta_i + U_{it} \tag{1}$$

Pero el estimador MCO no tiene en cuenta los "efectos fijos" θ_i, por lo que el modelo estimado es:

$$Y_{it} = \hat{\beta} + \hat{\alpha} D_i \tag{2}$$

Dónde el estimador MCO de α es igual a:

$$\hat{\alpha}_{MCO} = \frac{\sum_{i=1}^{n}(y_{it} - \bar{y}_t)(D_i - \bar{D})}{\sum_{i=1}^{n}(D_i - \bar{D})^2} \tag{3}$$

Reemplazando el valor de y_{it} por su expresión y el de $\bar{y}_t$ por $\beta + \alpha_i \bar{D}_i + \bar{\theta}_i + \bar{U}_t$ se obtiene la siguiente expresión para el estimador:

$$\hat{\alpha}_{MCO} = \alpha + \frac{\sum_{i=1}^{n}(\theta_t - \bar{\theta})(D_i - \bar{D})}{\sum_{i=1}^{n}(D_i - \bar{D})^2} + \frac{\sum_{i=1}^{n}(U_{it} - \bar{U}_t)(D_i - \bar{D})}{\sum_{i=1}^{n}(D_i - \bar{D})^2} \tag{4}$$

El primer término es el valor verdadero del parámetro α, el segundo es el cociente entre la covarianza entre θ_i y D_i y la varianza de D_i (ambas diferentes de cero) y el último término es igual a cero, por lo que el sesgo del estimador puede expresarse de la siguiente forma:

$$sesgo(\alpha) = \hat{\alpha}_{MCO} - \alpha = \frac{cov(\theta_i, D_i)}{var(D_i)} \tag{5}$$

El signo de la covarianza anterior debería ser negativo, dado lo expuesto anteriormente, por lo que el estimador de mínimos cuadrados ordinarios estará sesgado hacia abajo.

Pregunta 2

Siguiendo a Heckman et Al.[3] un instrumento cualquiera, denominado Z_i, siempre que:

1. Z_i esté correlacionado con al variable de interés D_i, es decir: $E(D_i Z_i) \neq 0$.
2. Z_i no esté correlacionada con ningún otro determinante de y_{it}, es decir, debe influenciar al resultado sólo por medio de su correlación con D_i.
3. los α_i son iguales para todos los individuos i, si no es así, debe requerirse que los individuos no basen su decisión de participar en él, sino será el estimador inconsistente para estimar $E(\alpha_i / D_i = 1)$.

En la simulación el instrumento es válido por construcción, ya que la variable se genera de manera independiente de todos los demás determinantes de y_{it}, salvo de D_i ya que la participación del programa depende de Z_i a través del costo de participar, definido como $c_i = Z_i \Phi + V_i$. Esto implica que el instrumento es válido en el caso de coeficientes fijos, mientras que no puede usarse con coeficientes aleatorios porque la decisión de participar depende de Z_i y el estimador será inconsistente en dicho caso.

Pregunta 3

Nuevamente siguiendo al artículo, el resultado puede expresarse en términos de resultados potenciales como:

$$y_{it} = D_i y_{1it} + (1 - D_i) y_{0it} \tag{6}$$

Aplicando el operador de esperanza condicional, al igual que Angrist y Pischke (1999)[1], pág. 14, se obtienen las siguientes expresiones:

$$E(y_{it}/D_i = 1) - E(y_{it}/D_i = 0) = \underbrace{E(y_{1i}/D_i = 1) - E(y_{0i}/D_i = 1)}_{\text{ATT}} + \underbrace{E(y_{0i}/D_i = 1) - E(y_{0i}/D_i = 0)}_{\text{Sesgo de selección}} \tag{7}$$

Reemplazando la expresión por la ecuación del ingreso del modelo simulado se obtiene:

$$\begin{aligned}
E(y_{it}/D_i = 1) &= \beta + E(\alpha_i/D_i = 1) + E(\theta_i/D_i = 1) + E(U_{it}/D_i = 1) \\
E(y_{it}/D_i = 0) &= \beta + E(\theta_i/D_i = 0) + E(U_{it}/D_i = 0) \\
E(y_{it}/D_i = 1) - E(y_{it}/D_i = 0) &= \underbrace{E(\alpha_i/D_i = 1)}_{\text{ATT}} + \underbrace{[(E(\theta_i/D_i = 1) - E(\theta_i/D_i = 0)) + (E(U_{it}/D_i = 1) - E(U_{it}/D_i = 0)]}_{\text{Sesgo de selección}}
\end{aligned}$$

Esto implica que para éstos datos simulados si existe sesgo de selección debido a que los individuos se autoseleccionan en el programa, lo cual se evidencia en la correlación existente entre D_i y los términos de error no observables θ_i y U_{it}.

Los sesgos pueden originarse en tres fuentes:

α_i : individuos que se autoseleccionan porque creen que obtendrán beneficios positivos del programa al sobreestimar el verdadero ATT.

θ_i : individuos que se seleccionan porque su costo de oportunidad en el momento de entrenarse es bajo ya que su ingreso es bajo y subestiman el verdadero ATT.

U_{it} : individuos que poseen ingresos bajos debido a que sus ingresos en el momento $t - 1$ también fueron bajos ya que U_{it} es un AR(1).

En el experimento de montecarlo se obtuvo como resultado un valor de 612,34 para $E(\alpha_i/D_i = 1)$ y de 100,2 para $E(\alpha_i)$ en promedio para las 100 muestras, lo que evidencia la selección en base a dicho parámetro para el caso de coeficientes comunes, cosa que no ocurre con coeficientes comunes, dónde ambos son idénticos por ser $\alpha = 100$.

Pregunta 4

Las respuestas de ésta pregunta y de los siguientes se basan en los resultados que se exponen en el Cuadro 1. En la columna (1) de dicho cuadro se muestra el sesgo del estimador MCO, en todos los casos el coeficiente posee signo negativo debido a que las personas con θ_i bajo tienen más incentivos a participar, éste sesgo por un lado persiste en el tiempo, ya que a media que avanzan los períodos continúa, pero se reduce progresivamente a medida que se aleja del período de implementación. La persistencia se debe a la existencia del parámetro ρ en el término de error autoregresivo, el cual hace que el sesgo persista en el tiempo, aunque cada vez con menor intensidad.

En la columna (2) se repiten las estimaciones, pero ahora el parámetro de interés es $E(\alpha)$, y se observa cómo el sesgo cambia de signo y de magnitud, siendo ahora fuertemente positivo porque las personas con valores altos de α_i sesgan hacia arriba el estimador en gran medida, la causa radica en la dependencia de D_i en α_i.

Pregunta 5

El estimador de diferencias en diferencias se obtiene de la siguiente forma: primero se reescribe el modelo utilizando una dummy adicional, la que se denominará d_t tal que:

$$d_t = \begin{cases} 1 & \text{Si t=1} \\ 0 & \text{Caso contrario} \end{cases}$$

De ésta forma, el modelo de ingresos puede expresarse de la siguiente forma:

$$y_{it} = \beta + \lambda d_t + \gamma D_i + \alpha_i (d_t \times D_i + U_{it} \tag{8}$$

Cuadro 1: Sesgo promedio de los estimadores del efecto tratamiento

	Caso base (Coef. aleatorios) Parámetro de interés: $E(\alpha/D=1)=612,34$ No apareado (1)	Caso base (Coef. Aleatorios) Parámetro de interés: $E(\alpha)=100,2$ No apareado (2)	Caso base (Coef. comunes) Parámetro de interés: $E(\alpha)=100$ No apareado (3)	Caso base (Coef. comunes) Parámetro de interés: $E(\alpha)=100$ Apareado (4)
Cross section k+1				
media	-144.7	367.4	-732.7	-440.6
sd	54.5	60.04	54.5	63.03
Cross section k+2				
media	-124.4	387.8	-630.6	-366.7
sd	52.44	57.33	52.44	69.38
Cross section k+3				
media	-111.8	400.4	-553.5	-315.3
sd	55.58	60.99	55.58	74.63
Cross section k+4				
media	-100.5	411.6	-487.9	-271.1
sd	51.68	55.44	51.68	74.57
DiD (-1,3)				
media	32.82	545	157.3	-16.58
sd	53.77	53.7	53.77	71.97
DiD (-3,3)				
media	-7.64	504.5	-52.09	-222
sd	60.68	63.26	60.68	84.59
DiD (-5,3)				
media	-45.08	467.1	-207.9	-210.6
sd	49.57	54.33	49.57	80.86
IV $k+1$				
media	-331.1	181	63.6	5.58
sd	1714	1713	1714	5816
IV $k+2$				
media	-72.46	439.7	65.27	17.04
sd	1850	1849	1850	4024
IV $k+3$				
media	-190.5	321.6	60.55	7.63
sd	2463	2463	2463	4549
IV $k+4$				
media	-256.8	255.3	66.11	15
sd	2846	2845	2846	7180

Entonces el efecto tratamiento en los tratados es igual a:

$$ATET = (E(y_{1it}/D_i = 1) - E(y_{1it}/D_i = 0)) - (E(y_{0it}/D_i = 1) - E(y_{0it}/D_i = 0)) = \alpha_i \tag{9}$$

El estimador se implementa realizando una regresión de la diferencia entre el valor de y_{it} para los valores de los períodos especificados de la siguiente forma:

$$y_{it} - y_{it-j} = \lambda + \alpha_i D_i + U_{it} \tag{10}$$

Y el coeficiente de la variable α_i es el que se reporta en el Cuadro 1, y j es el año base, del cual se consideran tres opciones: $k-1$, $k-3$ y $k-5$. Si el modelo estuviera bien especificado, las estimaciones de α_i tendrían que ser similares, pero se observa que el sesgo es positivo en cuando el período de referencia es $k-1$ y negativo para los períodos más lejanos, ésto se debe a que el ingreso del período de implementación es anormalmente bajo para quienes participan y la participación en éste, depende del ingreso en éste período.

El estimador DiD elimina el sesgo cuando se origina en θ_i pero no logra eliminarlo cuando se origina en U_{it}, por tal motivo las estimaciones continúan estando sesgadas, a pesar de ser éste inferior que para el caso del estimador de Mínimos cuadrados ordinarios.

En la columna (2) aparecen los resultados cuando el coeficiente de interés es $E(\alpha)$, la magnitud del sesgo también es mayor y de signo opuesto al de MCO, por lo que no se logra eliminar el sesgo remanente generado por la estructura del término de error.

El llamado .Ashenfelter's dip.es un fenómeno que ocurre cuando no se cumple el supuesto de identificación del estimador de diferencias en diferencias. El mencionado supuesto de identificación del estimador DiD afirma que, en ausencia del programa de entrenamineto, el cambio en los ingresos entre dos períodos de tiempo t y t' debería haber sido el mismo para los que participan como para los que no, esto es que se cumpla:

$$E(Y_{0t} - Y_{0t'}/D = 1) = E(Y_{0t} - Y_{0t'}/D = 0) \tag{11}$$

Ashenfelter (1978)[2] observó un hecho estilizado, el cual consiste en que, previo a inscribirse en un programa de entrenamiento los participantes experimentan una caída en sus ingresos, tanto en términos absolutos como relativo a los del grupo de control. Éste fenómeno sugiere que al menos una parte del incremento de los ingresos posterior a la implementación del programa se debe a una reversión del ingreso permanente que fuera interrumpido temporalmente por un shock adverso.

El supuesto de identificación del estimador de diferencias en diferencias puede no cumplirse en la medida que el momento base t' coincida con el momento del "dip"transitorio y, si los no participantes no experimentan la mencionada caída, el sendero temporal de los ingresos será diferente entre participantes y no participantes entre los momentos t y t', en este caso el estimador DiD sobreestimará el efecto del entrenamiento en los participantes.

Para evaluar la existencia del "dip.es necesario contar primero con los datos del ingreso promedio por cada período de cuatro grupos, donde w es la ponderación de la observación apareada.

- Participantes del programa (individuos con $D_i = 1$)
- No participantes del programa (individuos con $D_i = 0$)
- No participantes apareados (individuos con $D_i = 0$ y $w \neq 0$)
- No participantes apareados (individuos con $D_i = 0$ y $w = 0$)

En el primer gráfico se muestra el .Ashenfelter's dip"para el caso de coeficientes fijos, dónde claramente hay una caída en el ingreso de los individuos que participan del programa en el año previo a iniciarlo. Cabe señalara que la participación en el programa de éstos, si bien aumenta sus ingresos, no llega a igualar a los de los no participantes, debido a que el obtener un beneficio positivo por participar implica que los individuos poseen una productividad baja y por lo tanto inferior a la media de la población no participante. Para el caso de coeficientes aleatorios se observa que los resultados son muy diferentes, puesto que la figura del ingreso promedio para cada período difiere notablemente de la obtenida por Heckman et Al. (Fig. 15, pág. 2023) porque se observa un salto de gran magnitud en los ingresos del grupo tratado a partir del período 7, debido a la magnitud de α_i.

El gráfico de las distribuciones de los θ_i muestra que éstas difieren bastante entre participantes y no participantes en el caso de los coeficientes comunes.

Como contracara de lo anterior, las distribuciones de los distintos subgrupos en el caso de los coeficientes aleatorios son muy similares, reflejando con mucha exactitud los resultados de Heckman.

Resumiendo, la diferencia entre el caso de coeficientes fijos y el de coeficientes aleatorios reside en el mecanismo de seleccíon, porque en el segundo sólo dependerá de θ_i y de U_{it} ya que la ganancia de participar es igual para todos ya que α es idéntico para todos los individuos, y es ésta autoselección la que se refleja en el "dip"mas pronunciado para éste caso.

Pregunta 6

El estimador de variables instrumentales utiliza a Z_i, característica observable de la ecuación de participación como instrumento de D_i en la ecuación de ingresos y estimará consistentemente a $E(\alpha_i/D_i = 1)$ si se cumple que $E(Z_i D_i) \neq 0$, $E(Z_i \theta_i) = 0$ y $E(Z_i \epsilon_i) = 0$ y que el modelo sea el de coeficientes comunes, ya que la autoselección en el modelo de coeficientes aleatorios hace que el sesgo sea de gran magnitud y mucha variabilidad en tal caso.

En el caso de coeficientes comunes, el estimador de VI, si bien no elimina el sesgo, lo reduce considerablemente, de igual forma se reduce la variabilidad de las estimaciones del parámetro α.

La simulación de montecarlo guarda también los valores de la correlación entre D_i y Z_i para cada una de las 100 réplicas en los casos de coeficientes aleatorios y comunes, para el primer caso el valor promedio es de $-0{,}0000582$ y su valor p promedio es de $0,236$, por lo que la correlación no es estadísticamente significativa en este caso. Por otro lado, para los coeficientes aleatorios, el valor de la correlación es $-0,00027$ y su valor p es igual a 0, de modo que a pesa de ser muy bajo es significativo, tal cual debería suceder, ya que el instrumento es válido sólo para el segundo caso.

El coeficiente del estimador de variables instrumentales representa al efecto del tratamiento sobre quienes cambian de condición como respuesta a un cambio en Z_i, en este caso el instrumento es una característica personal que se utiliza para predecir a la participación, D_i, el coeficiente es, entonces:

$$\alpha_{VI} = \frac{cov(y_i, Z_i)}{cov(D_i, Z_i)} = \frac{\dfrac{cov(y_i, Z_i)}{var(Z_i)}}{\dfrac{cov(D_i, Z_i)}{var(Z_i)}} \tag{12}$$

Y se interpreta como el cambio marginal en la variable de resultado (ingreso) con relación a la característica personal Z_i para los individuos que modifican su situación ante cambios en D_i en respuesta a una variación muestral en dicha característica.

Pregunta 7

Las respuestas de ésta sección fueron ya analizadas en relación a los puntos anteriores para los tres tipos de estimadores estudiados.

Pregunta 8

En el último punto se implementa un estimador de apareamiento (o *matching*) por medio del método del vecino más cercano con reposición, el método realiza una comparación de los 100 individuos tratados en cada muestra con un subconjunto del grupo de control de manera tal que éstos individuos con los que se compara posean características similares a los del grupo de tratamiento, dónde por "similaridad" debe entenderse que cumplan un cierto criterio de distancia, siendo el criterio en cuestión el del vecino más cercano.

En la columna (4) se presentan los resultados que surgen de calcular los estimadores a la muestra apareada para el caso de coeficientes constantes. En el caso del estimador de mínimos cuadrados, el apareamiento no elimina el sesgo sino que lo balancea entre los diferentes grupos, el estimador MCO sigue estando sesgado hacia abajo, pero el sesgo se reduce notablemente si se lo compara con los resultados de la columna (3).

Para el estimador de Diferencias en diferencias ocurre algo similar, el sesgo se reduce notoriamente, pero no elimina su inconsistencia, ya que el origen se encuentra en la estructura de los errores y no en la presencia de θ_i únicamente, en cuyo casi si se lograría eliminar dicho sesgo, en cambio sólo se consiguen resultados mixtos para este estimador.

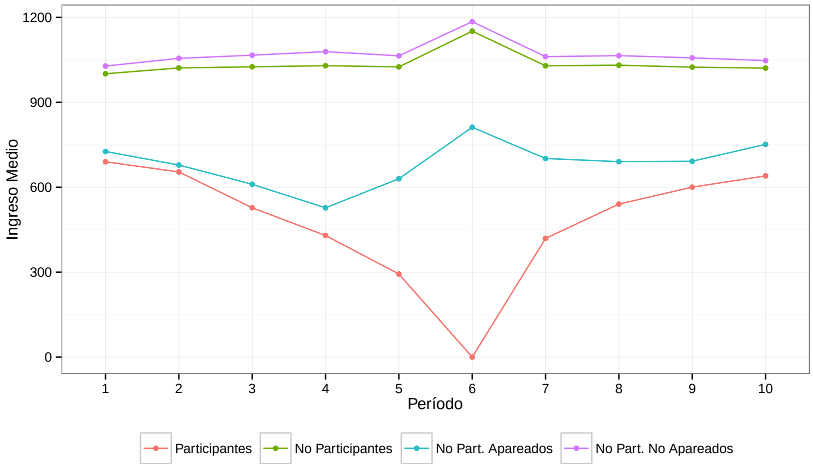
Por último, para el estimador de variables instrumentales, el sesgo es el menor de todos los casos, ya que, por un lado, es sólo válido para el estimador de coeficientes comunes (principal fuente de reducción del sesgo vis-à-vis el estimador MCO) y luego el balanceo de la muestra logra otra reducción en éste haciéndolo bastante pequeño, y cada vez mas bajo a medida que transcurren los períodos de tiempo, pero cómo contracara de éste, los errores estándar se "inflan" en gran medida, afectando la significatividad de los coeficientes estimados.

Referencias

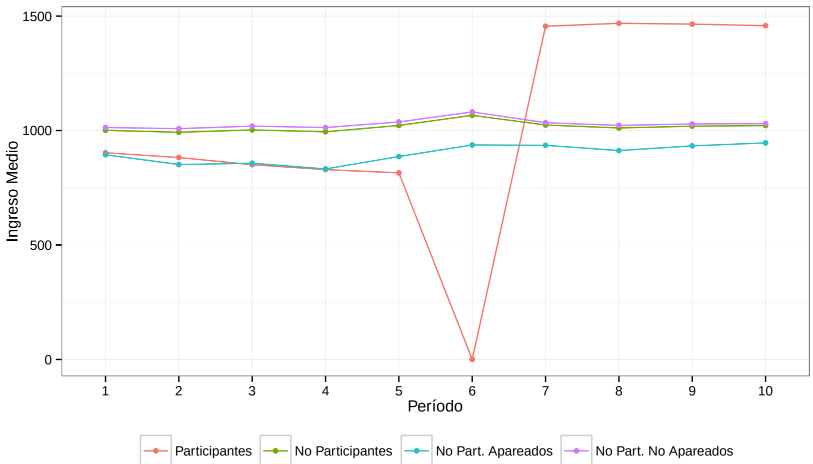
Angrist, Joshua y Jörn Steffen Pischke (2009). *Mostly Harmless Econometrics*. Princeton, NJ: Princeton University Press.

Ashenfelter, Orley (1978). “Estimating the Effect of Training Programs on Earnings”. En: *Review of Economics and Statistics* 6.1, págs. 47-57.

Heckman, James, Robert Lalonde y Jeffrey Smith (1999). “The Economics and Econometrics of Active Labor Market Programs”. En: *Handbook of Labor Economics*. Ed. por Orley Ashenfelter y David Card. Vol. 3A. Elsevier Science Publishers, págs. 1865-2097.

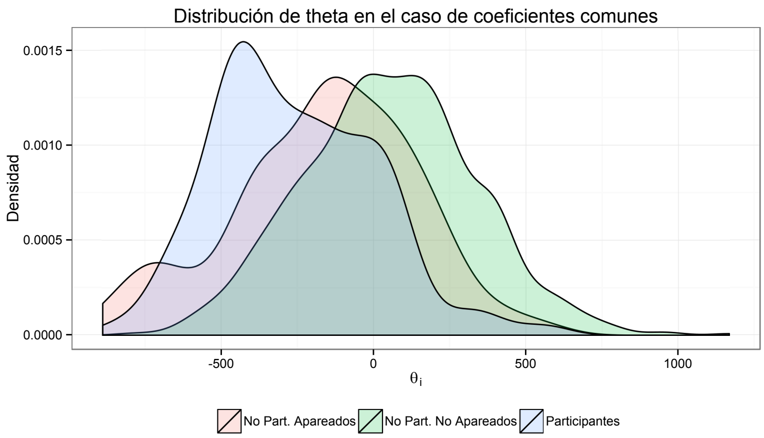


(a) Ingresos medios por período (Coeficientes Constantes)

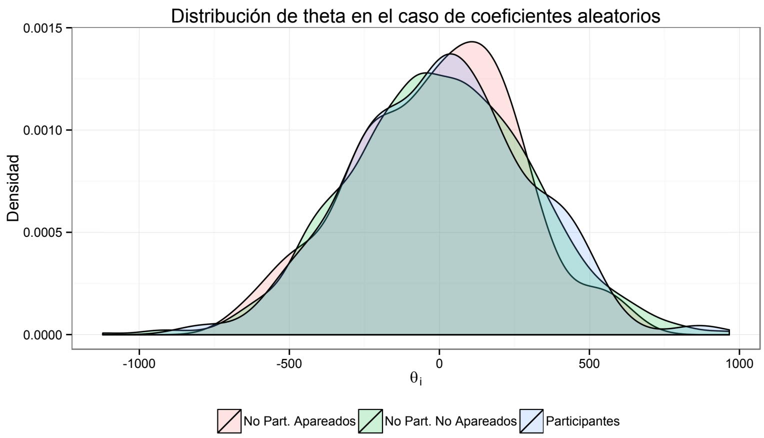


(b) Ingresos medios por período (Coeficientes Aleatorios)

Figura 1: *Ashenfelter's dip*



(a) Distribución de θ_i (Coeficientes Constantes)



(b) Distribución de θ_i (Coeficientes Aleatorios)

Figura 2: Distribución del término del efecto fijo para los distintos casos