Economía Aplicada

Estimador de diferencias en diferencias

Departamento de Economía
Universidad Carlos III de Madrid

Ver Wooldridge cap.13

Análisis de Política: Diferencias-en-Diferencias

- En muchos casos la variable de interés cambia en el tiempo para un grupo de individuos, una provincia, una cohorte. Por ejemplo, políticas estatales o provinciales sobre beneficios de desempleo para desempleados de larga duración pueden cambiar en el tiempo pero están fijas entre los trabajadores de esos estados.
- La fuente de sesgo de variable omitida en estos casos puede ser la presencia de variables no observables a nivel de estado y año.
- En estos casos podemos utilizar la estrategia de identificación llamada de Diferencias-en-Diferencias (DD o dif-en-dif).
- Para utilizar esta estrategia es posible usar secciones cruzadas repetidas.

Diferencias-en-Diferencias: intuición

- Supongamos que el tratamiento se asigna aleatoriamente a ciertas unidades (o al menos puede considerarse como si fuese aleatorio).
- Para estimar el efecto del tratamiento, podemos simplemente comparar las unidades tratadas antes y después del tratamiento.
- El problema es que podemos capturar los efectos de otros factores que hayan cambiado durante el tratamiento.
- Por lo tanto, usaremos un grupo de control para descontar estos posibles factores y aislar el efecto del tratamiento.

Diferencias-en-Diferencias: el primer ejemplo 1/2

- El primer ejemplo de aplicar DD se atribuye a Snow(1855). Este médico quería estudiar la hipótesis de que el cólera se transmitía por beber agua contaminada. En 1854, encontró un experimento natural útil para responder a esta pregunta.
- En 1849 dos compañías de agua, SVC y LC, obtenían el agua del contaminado Támesis en el centro de Londres. En 1852 la compañía LC comenzó a obtener su agua de una parte más limpia del río.
- Una comparación simple sería comparar las tasas de mortalidad en los distritos servidos por la compañía que cambió su fuente de agua, en 1852 vs 1849.

Diferencias-en-Diferencias: el primer ejemplo 2/2

- Pero esa comparación podría capturar los efectos de otros factores que pudieron haber cambiado en Londres en ese período (mejores medios para combatir el cólera, mayor información, etc.).
- Snow utilizó a los distritos servidos por la compañía que no hizo cambios como grupo de control.
- Snow comparó los cambios en las tasas de mortalidad por cólera entre 1854 y 1849 en distritos servidos por estas dos compañías.
- Encontró que las tasas de mortalidad en los distritos servidos por LC cayeron en comparación con el cambio producido en los distritos servidos por SVC.

Nueva Notación

- $Y(t)$: resultado observado en el período t
- $Y_1(t)$: resultado en el período t con tratamiento
- $Y_0(t)$: resultado en el período t sin tratamiento
- $t = 0$: antes del tratamiento
- $t = 1$: después del tratamiento

$$Y(t) = D * Y_1(t) + (1 - D) * Y_0(t)$$

Diferencias-en-Diferencias: un cálculo simple

- Se puede calcular la diferencia entre los dos grupos antes del tratamiento (D_1), la diferencia después del tratamiento (D_2) y luego calcular la "diferencia de las diferencias".

Resultados Promedio por Grupo y Período

	Tratamiento	Control	Diferencia
Antes	$E[Y_0(0)\|D=1]$	$E[Y_0(0)\|D=0]$	D_1
Después	$E[Y_1(1)\|D=1]$	$E[Y_0(1)\|D=0]$	D_2

- El estimador de DID es entonces: $DID = D_2 - D_1$
- Que equivale a:
$\{E[Y_1(1)|D=1] - E[Y_0(1)|D=0]\} - \{E[Y_0(0)|D=1] - E[Y_0(0)|D=0]\}$

Diferencias-en-Diferencias: un cálculo simple

- También se puede calcular el cambio observado para el grupo de tratamiento (D_T), el cambio para el grupo de control (D_C) y luego calcular la "diferencia de las diferencias".

Resultados Promedio por Grupo y Período

	Tratamiento	Control	Diferencia
Antes	$E[Y_0(0)\|D=1]$	$E[Y_0(0)\|D=0]$	D_1
Después	$E[Y_1(1)\|D=1]$	$E[Y_0(1)\|D=0]$	D_2
Diferencia	D_T	D_C	DID

- El estimador de DID es entonces: $DID = D_T - D_C$
- Que equivale a:
$\{E[Y_1(1)|D=1] - E[Y_0(0)|D=1]\} - \{E[Y_0(1)|D=0] - E[Y_0(0)|D=0]\}$

El cálculo en Snow

Snow: tasas de mortalidad cada 100,000 indiv.

	Distritos servidos por LC	Distritos servidos por SVC
1849	131	165
1853	60	114

Dos maneras:

$D_2 = (60 - 114)$ y $D_1 = (131 - 165)$: $DID = -54 - (-34) = -20$

$D_T = (60 - 131)$ y $D_C = (114 - 165)$: $DID = -71 - (-51) = -20$

Efecto en los tratados: α_{ATT}

¿Cómo afecta el tratamiento a los tratados?

- $\alpha_{ATT} = E\left[Y_1(1) - Y_0(1) | D = 1\right]$

- Problema: no observamos el resultado sin tratamiento para los tratados, en el período después del tratamiento: $E\left[Y_0(1) | D = 1\right]$.
- Idea: utilizar al grupo de control para obtener ese valor.

¿Qué información tenemos?

Tanto en secciones cruzadas repetidas como en datos de panel podemos estimar de manera consistente:

- el resultado promedio para los tratados sin tratamiento antes de que sean tratados: $E[Y_0(0)|D=1]$
- el resultado promedio para los no tratados sin tratamiento antes y después del tratamiento: $E[Y_0(1)-Y_0(0)|D=0]$, o sea el cambio experimentado por los no tratados.

El Supuesto de Caminos Paralelos

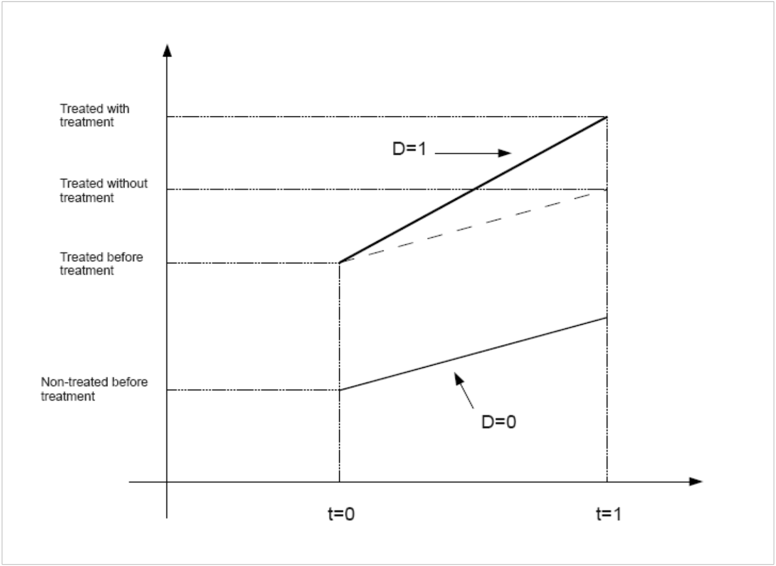
Caminos Paralelos

- El cambio en la variable de interés para los tratados hubiera sido el mismo que el cambio verificado para los no tratados si no hubiese habido cambio de política.

$$E\left[Y_0(1) - Y_0(0)|D = 1\right] = E\left[Y_0(1) - Y_0(0)|D = 0\right]$$

- Ejemplo: Las tasas de mortalidad en los distritos servidos por LC hubieran disminuido como lo hicieron en los distritos servidos por SVC, si LC no hubiese cambiado su agua. En ese caso hubiésemos esperado una disminución en las tasas de mortalidad de 51 en lugar de la observada: 71.

Una Interpretación Gráfica



Estimación de α_{ATT}

$$\alpha_{ATT} = E\left[Y_1(1) - Y_0(1)|D=1\right] = E\left[Y_1(1)|D=1\right] - E\left[Y_0(1)|D=1\right]$$

- Bajo el supuesto de caminos paralelos:

$$E\left[Y_0(1)|D=1\right] = E\left[Y_0(0)|D=1\right] + E\left[Y_0(1) - Y_0(0)|D=0\right]$$

- Entonces:

$$\begin{aligned}\alpha_{ATT} &= E\left[Y_1(1)|D=1\right] - E\left[Y_0(0)|D=1\right] - E\left[Y_0(1) - Y_0(0)|D=0\right] \\ &= E\left[Y_1(1) - Y_0(0)|D=1\right] - E\left[Y_0(1) - Y_0(0)|D=0\right]\end{aligned}$$

- Y observamos todos los promedios que necesitamos.

$$\hat{\alpha}_{ATT} = \left\{\overline{Y}_1(D=1) - \overline{Y}_0(D=1)\right\} - \left\{\overline{Y}_1(D=0) - \overline{Y}_0(D=0)\right\}$$

- Bajo condiciones generales, $\hat{\alpha}_{ATT}$ es consistente y asintóticamente normal.

MCO y el estimador de dif-en-dif (1/2)

- Podemos obtener el estimador en el contexto de regresiones de MCO:

$$Y_i = \beta_0 + \beta_1 D_i^{treated} + \beta_2 D_i^{after} + \beta_3 D_i^{treated} D_i^{after} + u_i$$

donde $D^{treated} = 1$ si la observación i pertenece al grupo de tratamiento y $D^{after} = 1$ si corresponde al período después del tratamiento.

- Entonces:
  - $E[Y_1(1)|D = 1] = E[Y|tratados, despues] = \beta_0 + \beta_1 + \beta_2 + \beta_3$
  - $E[Y_0(0)|D = 1] = E[Y|tratados, antes] = \beta_0 + \beta_1$
  - $E[Y_0(1)|D = 0] = E[Y|no\ tratados, despues] = \beta_0 + \beta_2$
  - $E[Y_0(0)|D = 0] = E[Y|no\ tratados, antes] = \beta_0$

MCO y el estimador de dif-en-dif (2/2)

- Volviendo a nuestra tabla:

Resultado Esperado por Grupo y Período

	Tratados	Control
Antes	$\beta_0 + \beta_1$	β_0
Después	$\beta_0 + \beta_1 + \beta_2 + \beta_3$	$\beta_0 + \beta_2$
Diferencia	$\beta_2 + \beta_3$	β_2

- Primera "dif": $E[Y_{despues} - Y_{antes} | tratados] = \beta_2 + \beta_3$
- Segunda "dif": $E[Y_{despues} - Y_{antes} | no\ tratados] = \beta_2$
- La "dif-en-dif": $(\beta_2 + \beta_3) - (\beta_2) = \beta_3$

El estimador de diferencias en diferencias condicional

$$Y_{it} = \beta_0 + \gamma X_i + \beta_1 D_i^{treated} + \beta_2 D_i^{after} + \beta_3 D_i^{treated} D_i^{after} + u_i$$

- El modelo se puede extender para incluir regresores adicionales: X_i representa a un conjunto de variables que pueden afectar a Y_i, por ejemplo características individuales antes del tratamiento.
- El estimador MCO de β_3 es el estimador de diferencias en diferencias condicional o con regresores adicionales.
- Se supone que el supuesto de caminos paralelos se cumple después de controlar por un conjunto de variables adicionales (condicional en X).
- Implementable tanto en datos de panel como en secciones cruzadas repetidas.

Ejemplo: El Efecto de las Leyes de Compensación a los Trabajadores en la Duración de las Bajas

- Meyer, Viscusi, y Durbin[1] estudiaron la duración de la compensación que reciben los trabajadores que sufren accidentes.
- En Julio de 1980, Kentucky incrementó el tope de los ingresos cubiertos por la compensación.
- Un incremento en el tope no afecta al beneficio de los trabajadores de bajos ingresos, pero hace menos costoso para un trabajador de ingresos altos acogerse a la compensación.
- Por ello el grupo de tratamiento son los trabajadores de altos ingresos y el grupo de control los trabajadores de bajos ingresos.

[1] Workers' Compensation and Injury Duration: Evidence from a Natural Experiment: AER (1995)

Ejemplo (cont.)

Con datos antes y después del incremento en el tope, por ej. 1979 y 1981

$$log(durat)_i = \beta_0 + \beta_1 highearn_i + \beta_2 D_i^{1981} + \beta_3 highearn_i D_i^{1981} + u_i$$

donde *durat* es la duración de la baja, *highearn* una variable binaria para trabajadores de altos ingresos y D^{1981} una variable que vale uno en 1981 y cero en 1979.

- β_1 captura la diferencia en duración promedio para trabajadores de altos y bajos ingresos antes del cambio.
- β_2 captura cambios en la duración promedio para los trabajadores de bajos ingresos.
- β_3 es el estimador de dif-en-dif: estima consistentemente el efecto de la política si los trabajadores de altos y bajos ingresos no sufrieron otros cambios diferenciales, solamente el incremento del tope.