

Semana 10. Diferencias en Diferencias

Equipo Econometría Avanzada

Universidad de los Andes

21 de octubre de 2022



Contenido

- 1 Contexto y Pregunta de investigación
- 2 Aproximación empírica

Contenido

1 Contexto y Pregunta de investigación

2 Aproximación empírica

Contexto y Pregunta de investigación

Lim & Sabrowski (2012) utilizan un panel de datos que contiene información sobre las principales firmas exportadoras en Turquía entre 2006 y 2015 para responder la siguiente pregunta:

¿Cuál es el impacto de que aliados comerciales sufran choques adversos sobre la diversificación de mercados en las firmas exportadoras?

Hechos:

- Siria era uno de los principales mercados de exportación de las firmas turcas.
- Debido a la crisis política de Siria, las importaciones del país cayeron drásticamente desde el tercer trimestre del 2011 ($t = 2011q3$).

Aproximación empírica

Comparación ingenua

Para evaluar el impacto del choque, necesitamos un grupo de tratamiento y uno de control. Se define el grupo tratamiento como aquellas firmas con una fuerte conexión con Siria antes de la guerra. En particular, los autores toman:

$$D_i = \text{ChoqueSiria}_i = \mathbb{1} \left[\frac{\text{Exportaciones a Siria}_i}{\text{Exportaciones totales}_i} > 0,5 \right]$$

Es tentador entonces para responder la pregunta usar el estimador:

$$\hat{\tau}_{CI} = \bar{Y}_{D=1} - \bar{Y}_{D=0} \text{ usando } t \geq 2011q3$$

Donde $\bar{Y}_{D=j}$ es el promedio de la variable Y para las firmas i con $D_i = j$

- ¿Es $\hat{\tau}_{CI}$ un estimador consistente del efecto de interés?

Aproximación empírica

Diferencias en Diferencias: 2x2

Hipótesis: Cuando exportar a Siria es muy importante para una firma, haber perdido ese mercado le obliga a diversificar los mercados a los que exporta. Por tanto, estar expuesto al choque debería generar una diversificación mayor.

Metodología: Se compara el numero de países a los que exportan las firmas “expuestas” vs. las “no expuestas”, antes y después del choque político en Siria. Esto se conoce como el *Método de Diferencias en Diferencias*.

Modelo estático

$$\begin{aligned} \text{Num_countries}_{it} = & \beta_0 + \beta_1 \text{Post}_t + \beta_2 \text{ChoqueSiria}_i \\ & + \tau \text{Post}_t \times \text{ChoqueSiria}_i + e_{it} \end{aligned}$$

Aproximación empírica

Supuestos de identificación:

1) Tendencias paralelas:

$$\mathbb{E}[Y_{i1}^0 - Y_{i0}^0 | D_i = 1] = \mathbb{E}[Y_{i1}^0 - Y_{i0}^0 | D_i = 0]$$

Intuición: 1) En ausencia del tratamiento, el cambio esperado en la variable dependiente del grupo de tratamiento es igual al cambio observado en esa variable para el grupo de control

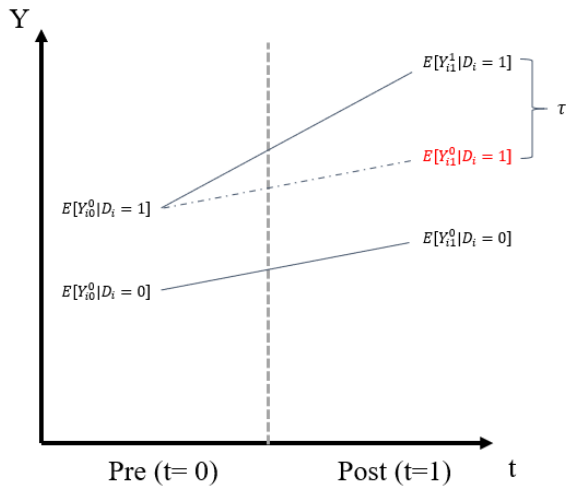
2) En ausencia del choque/programa, el grupo de tratamiento hubiera preservado en el tiempo la diferencia inicial en el periodo pre-tratamiento:

$$\mathbb{E}[Y_{i1}^0 | D_i = 1] = \underbrace{\mathbb{E}[Y_{i1}^0 | D_i = 0]}_{\text{Nivel del grupo de control en el post}} + \underbrace{\left(\mathbb{E}[Y_{i0}^0 | D_i = 1] - \mathbb{E}[Y_{i0}^0 | D_i = 0] \right)}_{\text{Diferencia inicial en el pre}}$$

2) No anticipación: $Y_{i0}^1 = Y_{i0}^0$ para todos los i tales que $D_i = 1$.

Intuición: Los valores de la variable independiente observados antes del momento de tratamiento no dependen de si voy a ser o no tratado en un futuro.

Aproximación empírica



Aproximación empírica

Diferencias en Diferencias: TWFE estático

Modelo estático con EF (TWFE)

$$\text{Num_countries}_{it} = \alpha_i + \theta_t + \tau \text{Post}_t \times \text{ChoqueSiria}_i + e_{it} \quad (1)$$

- θ_t es una especificación más flexible que Post_t .
- α_i es una especificación más flexible que ChoqueSiria_i .

Supongamos observamos también la industria j de la firma i . Podemos estimar alternativamente

$$\text{Num_countries}_{ijt} = \alpha_i + \theta_{jt} + \tau \text{Post}_t \times \text{ChoqueSiria}_i + e_{ijt} \quad (2)$$

- ¿Cuál es la diferencia entre estos modelos? ¿Alguno es siempre preferible?

Aproximación empírica

Diferencias en Diferencias: Modelo Dinámico

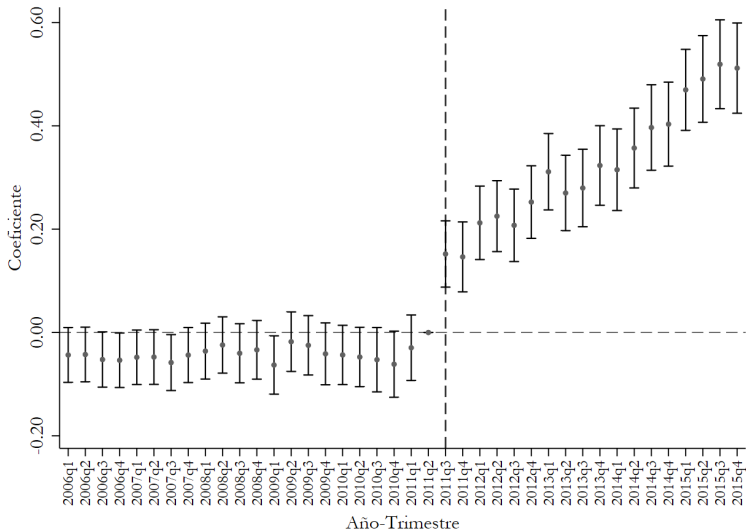
Modelo Dinámico:

$$Num_Countries_{ijt} = \alpha_i + \theta_{jt} + \sum_{\substack{k=2006q1 \\ k \neq 2011q2}}^{2015q4} \tau_k \mathbb{1}[t = k] \times Choquesiria_i + e_{ijt}$$

- No se pueden incluir todas las interacciones (multicolinealidad perfecta con α_i). Se suele omitir el periodo justo antes del tratamiento (en nuestro caso 2011q2). La interpretación de τ_k es con respecto al periodo omitido.
- Los rezagos (“lags”) permiten analizar efectos dinámicos del tratamiento (τ_k para $k \geq 2011q3$).
- Los efectos anticipatorios (“leads”) permiten validar el supuesto de identificación (τ_k para $k < 2011q3$).

Aproximación empírica

Leads and Lags.



Otros temas DiD

Corte Transversal Repetido

- **Corte transversal repetido:** No tenemos un panel de datos, pero observamos dos muestras **distintas** de la misma población. Es decir, no observamos exactamente a los mismos individuos en diferentes periodos de tiempo.
 - ▶ Gran Encuesta Integrada de Hogares.
 - ▶ Encuesta Nacional de Demografía y Salud
 - ▶ Saber 3ro, 5to, 9no y 11
- **Supuesto Adicional:**

Los individuos en la base de datos de corte transversal inicial pueden ser utilizados como sustitutos de los individuos de los grupos de tratamiento y control del corte transversal posterior.

 - ▶ Se necesita que las dos muestras sean aleatorias y representativas de la misma población.
- Aproximación por un modelo de regresión:

$$Y_{igt} = \alpha_g + \delta_t + \tau (\text{Post}_t \times D_i) + \theta X_i + u_{igt}$$

Otros temas DiD

Controles que varían entre unidades y en el tiempo

Es una práctica usual agregar controles adicionales para tratar de suavizar el supuesto de tendencias paralelas a tendencias paralelas condicionales. Al hacerlo, el modelo toma la forma de

$$Y_{it} = \alpha_i + \lambda_t + \tau (\text{Post}_t \times D_i) + \theta X_{it} + \epsilon_{it}$$

Pero ojo: ¡Incluir controles adicionales cambia los supuestos de identificación! También aplica para el modelo dinámico!

Otros temas DiD

Supuestos de identificación con controles adicionales

a) Tendencias paralelas condicionales:

$$\mathbb{E}[Y_{it}^0 - Y_{it}^1 | X, D_i = 1] = \mathbb{E}[Y_{it}^0 - Y_{it}^1 | X, D_i = 0]$$

Intuición: Condicional en tener los mismos X , las unidades de tratamiento y control hubieran evolucionado de manera paralela.

b) Soporte común: Existe $\varepsilon > 0$ tal que para todo $x \in \text{supp}(X)$:

$$\mathbb{P}(D_i = 1 | X = x) \geq \varepsilon$$

Intuición: Para cada valor de X existe al menos un tratado y un control con probabilidad positiva.

Otros temas DiD

Supuestos de identificación con controles adicionales

c) Efectos homogéneos en X :

$$ATT(X) = \mathbb{E}[Y_{it}^1 - Y_{it}^0 | X, D_i = 1] = \mathbb{E}[Y_{it}^1 - Y_{it}^0 | D_i = 1] = ATT$$

Intuición: El tratamiento afecta a todas las unidades tratadas por igual, independiente de su X .

d) No hay tendencias específicas en los grupos que dependen de X :

$$\mathbb{E}[Y_{i1} - Y_{i0} | D_i = d, X] = \mathbb{E}[Y_{i1} - Y_{i0} | D_i = d]; \quad d \in \{0, 1\}$$

Intuición: La evolución observada de la variable dependiente en el tiempo, tanto para el grupo de control como para el de tratamiento, es independiente de sus características X .

Conclusiones

- El modelo de diferencias en diferencias permite estimar efectos causales aun cuando existen diferencias sistemáticas entre los grupos que queremos comparar.
- **Supuesto de identificación:** Tendencias paralelas+No efectos anticipatorios.
 - ▶ Los individuos tratados se habrían comportado/habrían evolucionado de la misma forma que los no tratados en ausencia del tratamiento.
- Podemos verificar parcialmente el cumplimiento de tendencias paralelas cuando tenemos más de un periodo pre-tratamiento.
- En presencia de cortes transversales repetidos debemos adicionalmente suponer que los individuos de un corte transversal son sustitutos de los individuos de otro corte.
- Si queremos añadir controles adicionales y estimar por MCO, cambian los supuestos de identificación. Una alternativa es usar metodos no-paramétricos que sólo requieren el supuesto de tendencias paralelas condicionales y soporte común como en Sant'Anna & Zhao (2020).

¡Gracias!