

Visualización, Identificación y Estimación en el Diseño de Estudio de Eventos de Panel Lineal

Simon Freyaldenhoven¹ Christian
Hansen² Jorge Pérez Pérez³ Jesse M.
Shapiro⁴

¹Federal Reserve Bank of Philadelphia

²University of Chicago

³Banco de México

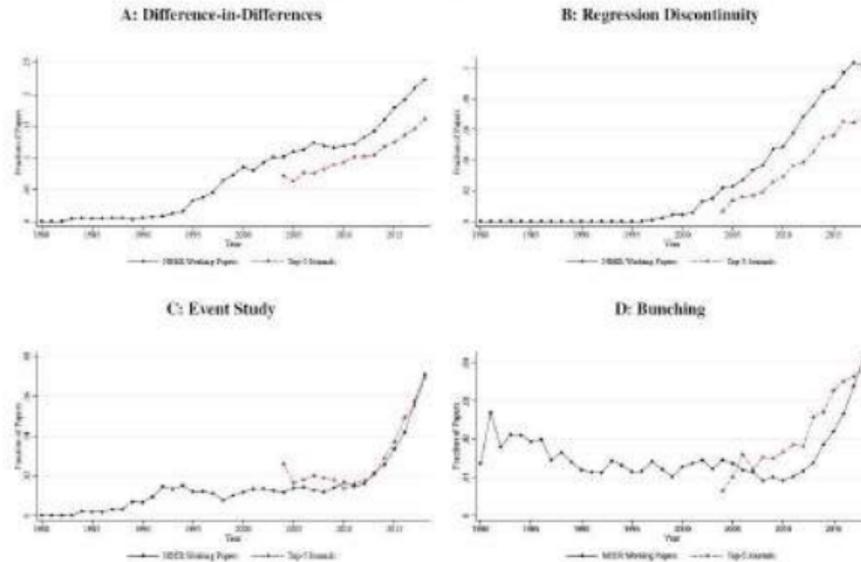
⁴Harvard University y NBER

Conferencia de Stata
Latinoamérica Sep 8, 2022

Motivación

- ▶ Las estimaciones de estudios de evento en panel se están volviendo cada vez más populares en economía

Figure IV: Quasi-Experimental Methods



Notes: This figure shows the fraction of papers referring to each type of quasi-experimental approach. See Table A.1 for a list of terms. The series show 5-year moving averages.

Fuente: [Currie et al. \(2020\)](#)

En esta presentación

1. Sugerencias para construir gráficos de estudio de evento
 - ▶ Paquete de Stata `xtevent` para facilitar la adopción de las sugerencias
2. Enfoques de identificación y su contenido económico
3. Desempeño de diferentes estimadores para diferentes procesos generadores de datos con contenido económico

Contexto

Modelo de panel lineal

Queremos estimar

- ▶ El efecto dinámico de una variable escalar de política z_{it}
- ▶ en una variable de resultado y_{it}
- ▶ con un panel observacional de unidades $i \in \{1, \dots, N\}$
- ▶ observado en una secuencia de tiempo $t \in \{1, \dots, T\}$

Ejemplos:

- ▶ Participación en un programa de entrenamiento laboral (e.g., [Ashenfelter 1978](#))
 - ▶ Las unidades i son individuos
 - ▶ z_{it} es una indicadora de participación en el programa
 - ▶ y_{it} es el ingreso laboral
- ▶ Efecto del salario mínimo (e.g., [Brown 1999](#))
 - ▶ Las unidades i son estados en EEUU
 - ▶ z_{it} es el nivel del salario mínimo
 - ▶ y_{it} es la tasa de desempleo de un grupo demográfico particular

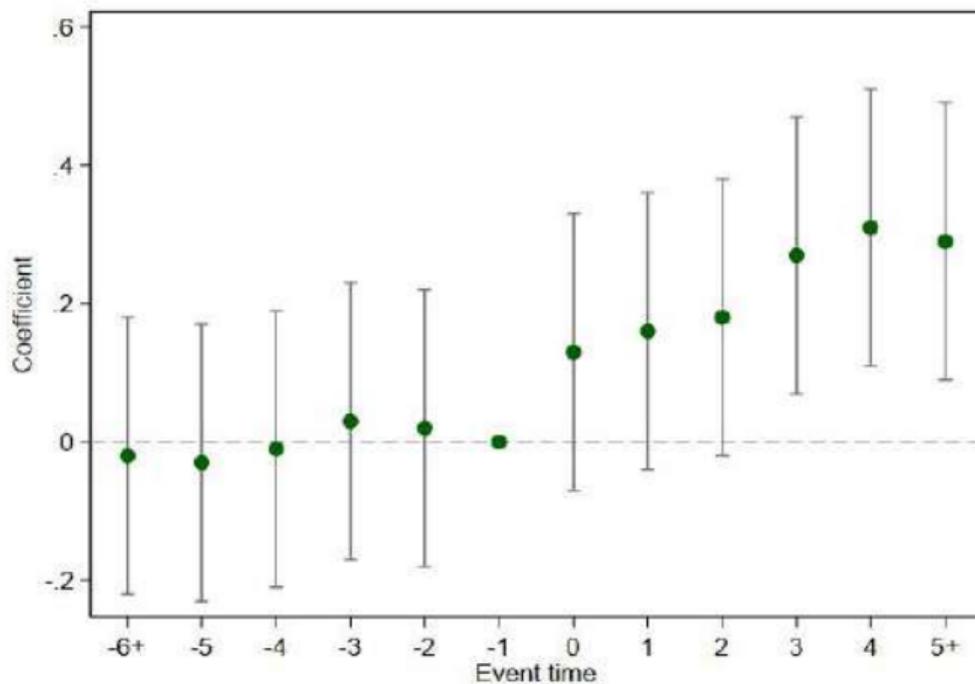
Modelo de Panel Lineal

$$y_{it} = \alpha_j + \gamma_t + \mathbf{q}'_{it}\psi + \sum_{m=-G}^M \beta_m z_{i,t-m} + C_{it} + \varepsilon_{it} \quad (\text{Modelo de Panel Lineal})$$

- ▶ α_j son efectos fijos de unidad y γ_t son efectos fijos de tiempo
- ▶ \mathbf{q}_{it} es un vector de controles
- ▶ C_{it} es una variable (potencialmente no observada) que puede estar correlacionada con la variable de política
- ▶ ε_{it} representa choques no observados que no están correlacionados con la variable de política
- ▶ Los parámetros de interés son $\{\beta_m\}_{m=-G}^M$
 - ▶ No hay efecto (directo) de la política más de G períodos en el pasado o más de M períodos en el futuro

Gráficos de estudio de evento

Gráfico típico de estudio de evento



Construcción del gráfico

$$y_{it} = \alpha_i + \gamma_t + \mathbf{d}'_{it}\psi + \sum_{m=-G}^M \beta_m z_{i,t-m} + C_{it} + \varepsilon_{it} \quad (\text{Modelo de Panel Lineal})$$

Para el gráfico queremos:

- ▶ Mostrar efectos acumulados de la política. Para eso cambiamos z_{it} por δz_{it}
- ▶ Mostrar dinámica antes de G y después de M . Para eso agregamos L_G adelantos adicionales y L_M rezagos adicionales

Ecuación a estimar

$$y_{it} = \sum_{k=-G-L_G}^{M+L_M-1} \delta_k \Delta Z_{i,t-k} + \delta_{M+L_M} Z_{i,t-M-L_M} + \delta_{-G-L_G-1} (-Z_{i,t+G+L_G+1}) \\ + \alpha_i + \gamma_t + \mathbf{q}'_{it} \boldsymbol{\psi} + C_{it} + \varepsilon_{it},$$

(Ecuación a estimar)

- ▶ Llamamos a k el *tiempo de evento*
- ▶ Llamamos a δ la *trayectoria en tiempo de evento* del resultado

Interpretación con adopción escalonada

$$\dots \sum_{k=-G-L_G}^{M+L_M-1} \delta_k \Delta z_{i,t-k} + \delta_{M+L_M} z_{i,t-M-L_M} + \delta_{-G-L_G-1} (-z_{i,t+G+L_G+1}) \dots$$

(Parte principal de la ecuación a estimar)

Adopción escalonada: Para cada unidad i , z_{it} arranca en cero y cambia a 1 en algún momento.

- ▶ $\Delta z_{i,t+k}$ indica si la unidad i adoptó la política exactamente k periodos después del periodo t ,
- ▶ $z_{i,t+G+L_G+1}$ indica si la unidad i habrá adoptado la política $G + L_G + 1$ periodos después del periodo t
- ▶ $z_{i,t-M-L_M}$ indica si la unidad i adoptó la política al menos $M + L_M$ periodos antes del periodo t

Interpretación de efectos acumulados de la política

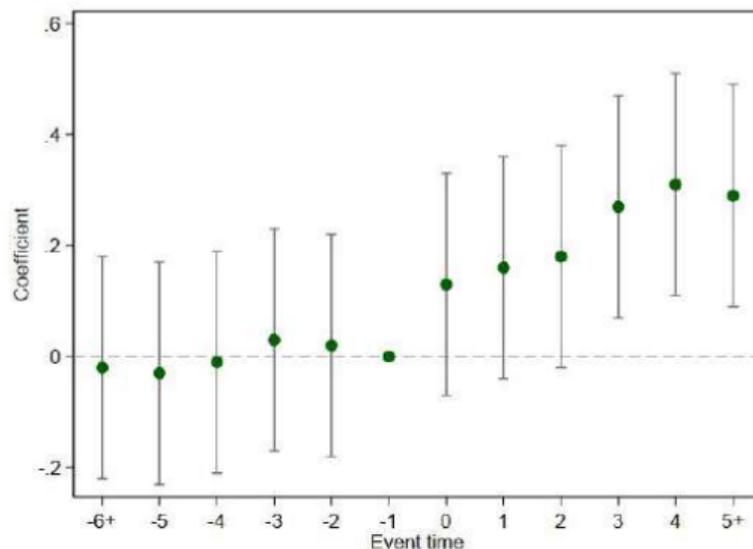
$$\dots \sum_{k=-G-L_G}^{M+L_M-1} \delta_k \Delta Z_{i,t-k} + \delta_{M+L_M} Z_{i,t-M-L_M} + \delta_{-G-L_G-1} (-Z_{i,t+G+L_G+1}) \dots$$

(Parte principal de la ecuación a estimar)

Los parámetros $\{\delta_k\}_{k=-G-L_G-1}^{k=M+L_M}$ pueden interpretarse como efectos acumulados de la política:

$$\delta_k = \begin{cases} 0 & \text{for } k < -G \\ \sum_{m=-G}^k \beta_m & \text{para } -G \leq k \leq M \\ \sum_{m=-G}^M \beta_m & \text{para } k > M. \end{cases} \quad (1)$$

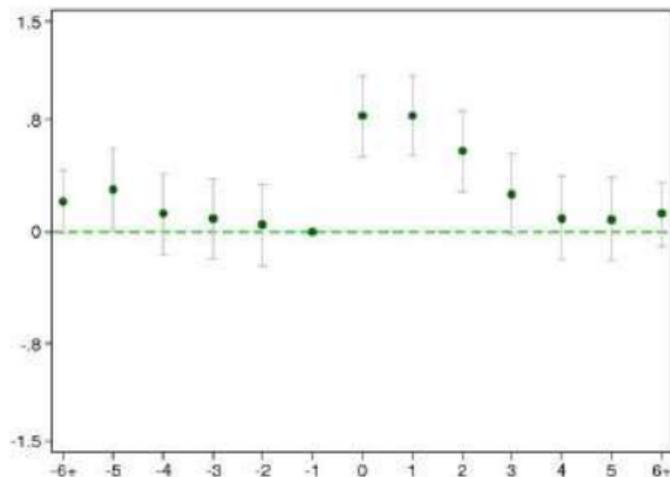
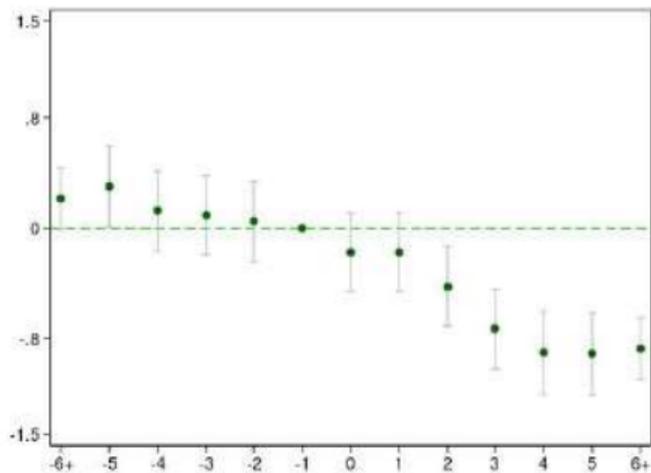
Definición de gráfico de estudio de evento



Los puntos en el gráfico son $\{(k, \hat{\delta}_k)\}_{k=-G-L_G-1}^{k=M+L_M}$.

Sugerencias

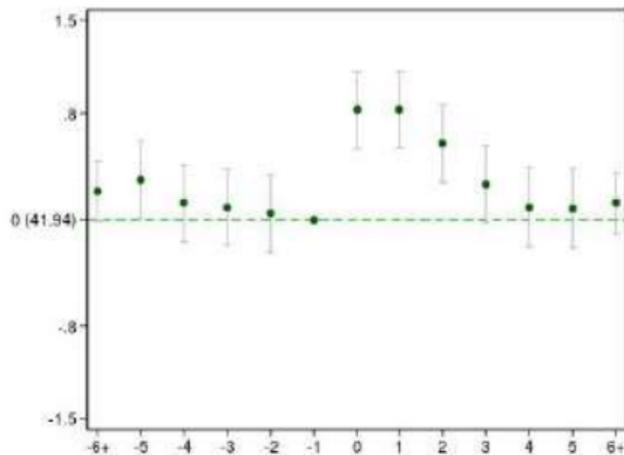
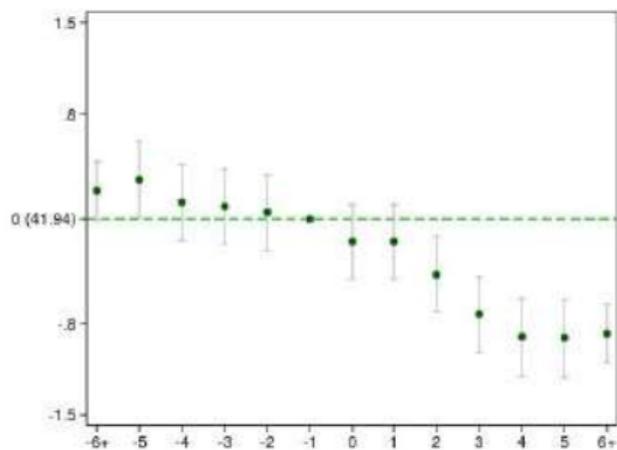
Normalización



Sugerencia 1

Normalizar $\delta_{-G-1} = 0$ en la ecuación a estimar. (En el ejemplo, $G = 0$)

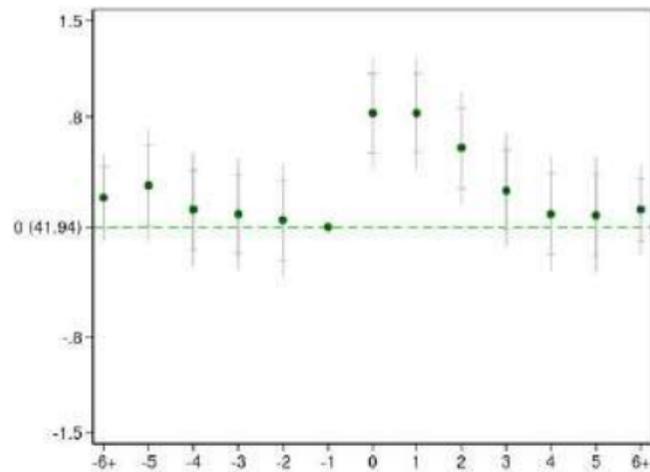
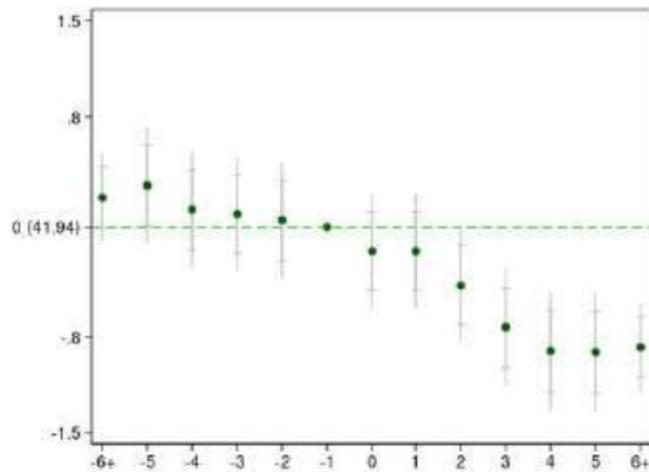
Magnitud



Sugerencia 2

Incluir un valor entre paréntesis que muestre el valor del promedio de la variable dependiente en el periodo correspondiente al coeficiente normalizado.

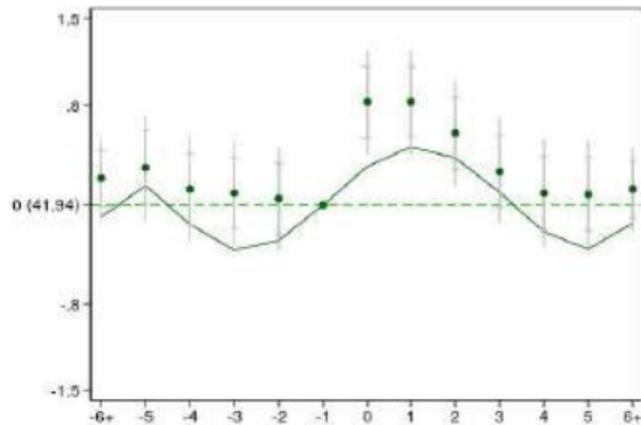
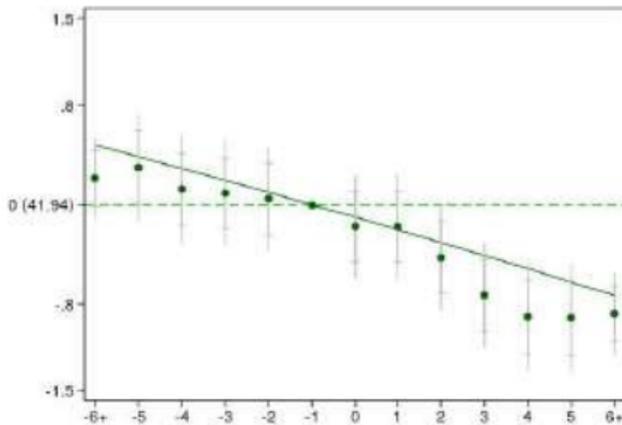
Inferencia



Sugerencia 3

Graficar intervalos de confianza uniformes además de los intervalos de confianza para cada uno de los elementos de la trayectoria.

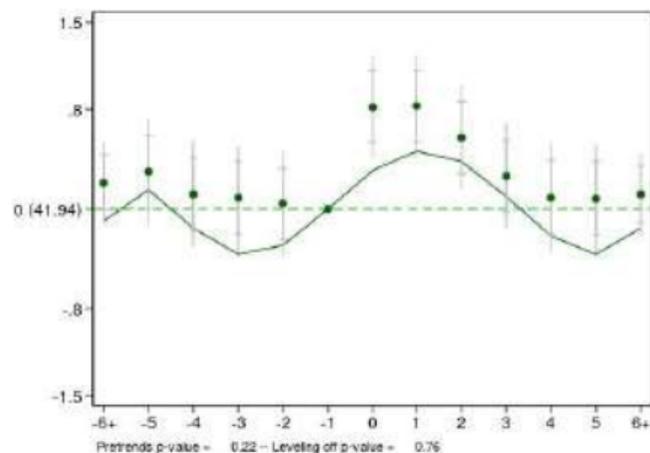
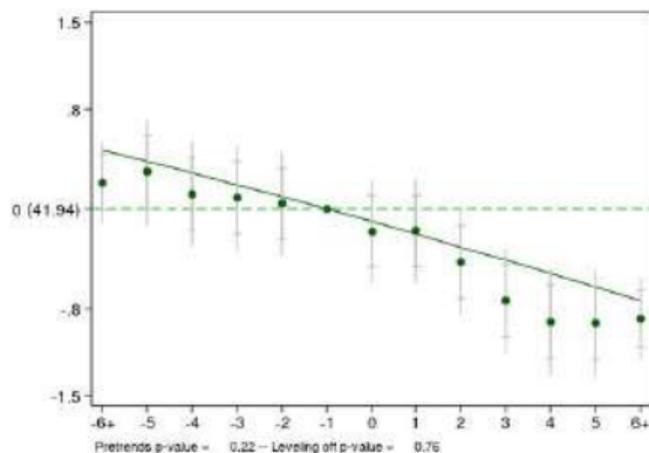
Trayectoria de variable omitida



Sugerencia 4

Mostrar la trayectoria más “suave” de una variable omitida que sería consistente con los datos. Específicamente, mostrar el polinomio que tenga el coeficiente más pequeño asociado al término de mayor orden, dentro de los polinomios de mínimo orden que pasan por la región de confianza de δ .

Pruebas de sobreidentificación



Sugerencia 5

Incluir en la leyenda p-valores para pruebas de Wald de las siguientes hipótesis;

$$H_0 : \delta_k = 0,$$

$$H_0 : \delta_M = \delta_{M+k},$$

$$-(G + L_G) \leq k < -G$$

$$0 < k \leq L_M$$

(no hay tendencias previas)

(los efectos se nivelan)

Implementando las sugerencias con `xtevent` en Stata

- ▶ Estudio de evento básico con adopción escalonada
`xtevent` y `eta` , `panelvar(i)` `timevar(t)`
`policyvar(z)` `window(5)` `impute(stag)`
- ▶ Gráfico de estudio de evento
`xteventplot`
- ▶ Trayectoria de efecto de variables omitidas
`xteventplot, smpath(line)`

Enfoques de identificación

Variables omitidas

$$y_{it} = \alpha_i + \gamma_t + \mathbf{q}'_{it}\psi + \sum_{m=-G}^M \beta_m z_{i,t-m} + C_{it} + \varepsilon_{it} \quad (\text{Modelo de Panel Lineal})$$

- ▶ En general C_{it} no se observa, así que la identificación de $\{\beta_m\}_{m=-G}^M$ requiere supuestos acerca de C_{it}
- ▶ En el documento se señalan muchos enfoques. En esta presentación solo hablamos de algunos de ellos

La variable omitida es de dimensión reducida

Supuesto 1

$$C_{it} = \lambda_i' F_t, \text{ y además:}$$

- a. $F_t = 0$ para todo t
 - ▶ “Todos los choques macro afectan a las unidades de la misma manera”
 - ▶ Estimar β con efectos fijos bidireccionales y controles
- b. $F_t = f(t)$ for $f(\cdot)$ para un conjunto conocido de funciones base de dimensión reducida
 - ▶ Aproximar posibles fuentes de sesgo con una tendencia
 - ▶ Estimar con efectos fijos bidireccionales controlando por tendencias de tiempo por unidad
- c. La dimensión de F_t es pequeña
 - ▶ Cada unidad tiene una respuesta única a choques comunes
 - ▶ Estimar con efectos fijos interactivos, efectos correlacionados comunes, o controles sintéticos

La variable omitida puede extrapolarse con información previa al evento

Supuesto 2

$$\mathbb{E}[C_{it}|z_i, \alpha_i, \gamma_t, \mathbf{q}_i] = \tilde{\alpha}_i + \tilde{\gamma}_t + \mathbf{q}'_{it}\tilde{\psi} + \sum_m \phi' f(m)z_{i,t-m} \quad (2)$$

para un conjunto conocido de funciones base de dimensión reducida $f(\cdot)$, y parámetros desconocidos $\tilde{\alpha}_i$, $\tilde{\gamma}_t$, $\tilde{\psi}$, y ϕ

Extrapolación de tendencias usando `xtevent`

- ▶ Estudio de evento básico y gráfico

```
xtevent y eta , panelvar(i) timevar(t)  
policyvar(z) window(5) impute(stag) plot
```

- ▶ Superponer la tendencia

```
xtevent y , panelvar(i) timevar(t) policyvar(z)  
window(5) trend(-3)  
xteventplot, overlay(trend)
```

- ▶ Restar la tendencia extrapolada

```
xteventplot
```

Proxy con ruido que no está relacionado con la política

Supuesto 3

Observamos un vector x_{it} que satisface

$$x_{it} = \alpha_j^x + \gamma_t^x + \psi^x \mathbf{q}_{it} + \Xi^x \eta_{it} + u_{it}, \quad (3)$$

El inobservable u_{it} satisface

$$\mathbb{E}[u_{it} | z_i, \alpha_j^x, \gamma_t^x, \mathbf{q}_i] = 0,$$

y la proyección de η_{it} en $\{z_{i,t-m}\}_{m=-G-L_G}^{M+L_M}$, \mathbf{q}_{it} e indicadores de unidad y tiempo, en la población, tiene al menos un coeficiente diferente de cero en $z_{i,t+m}$ para algún $m > G$

Proxy con ruido y estrategia de variables instrumentales en Stata con `xtevent`

- ▶ Estudio de evento para el resultado
`xtevent y , panelvar(i) timevar(t) policyvar(z)`
`window(5) proxy(x)`
- ▶ Gráfico de estudio de evento para el resultado
`xteventplot, y`
- ▶ Gráfico de estudio de evento para la proxy
`xteventplot, proxy`
- ▶ Alinear la proxy y el resultado
`xteventplot, overlay(iv)`
- ▶ Restar la variable omitida del resultado
`xteventplot`

Heterogeneidad

- ▶ Literatura reciente permite que los efectos de la política sean diferentes entre unidades
- ▶ La ecuación (**Modelo de Panel Lineal**) puede generalizarse para permitir diferencias entre cohortes de tratamiento $t^*(i)$

$$y_{it} = \alpha_i + \gamma_t + \mathbf{q}'_{it}\psi + \sum_{m=-G}^M \beta_{m,t^*(i)} z_{i,t-m} + C_{it} + \varepsilon_{it}.$$

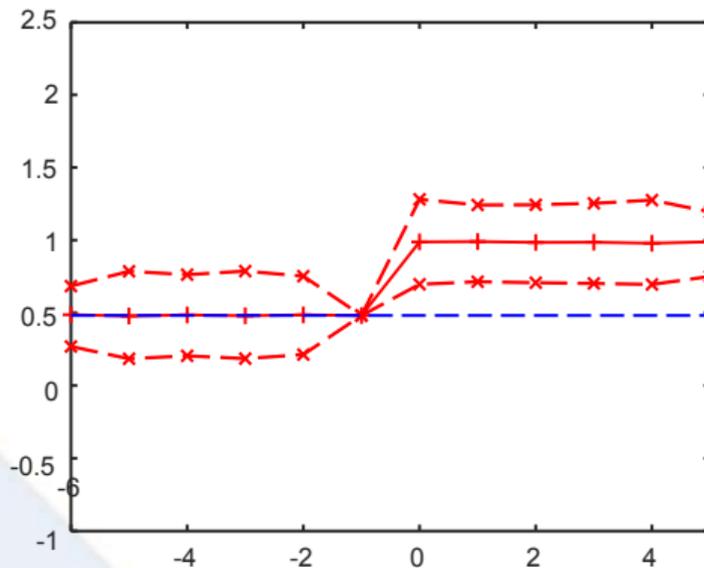
y luego pueden usarse los mismos enfoques de identificación

Simulaciones

~ Diseño de las simulaciones

- ▶ $N = 50, T = 40, \beta = 0,5$
- ▶ Diferentes dinámicas para la variable omitida, adopción endógena de política

Figura: Resultado sin variable omitida $y_{it} - C_{it}$



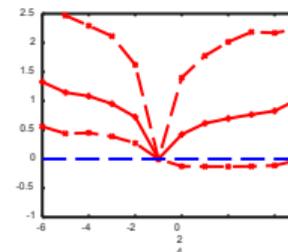
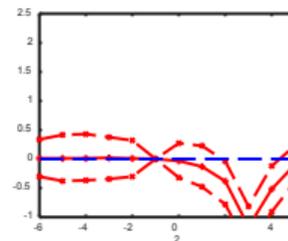
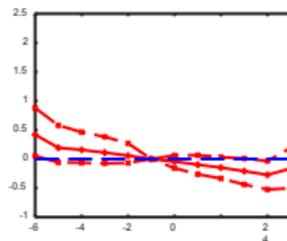
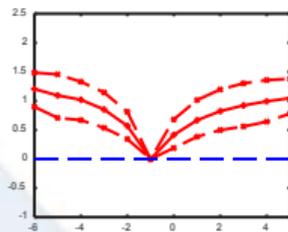
Resumen de los procesos generadores de datos

Mean-rev. trend

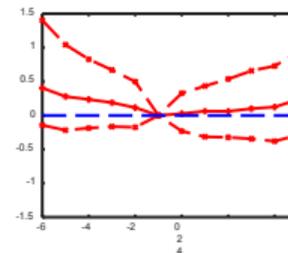
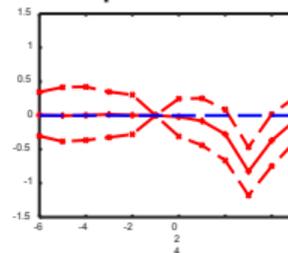
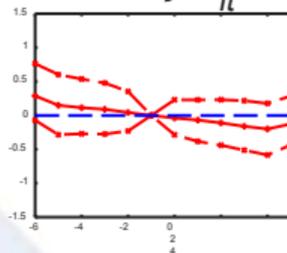
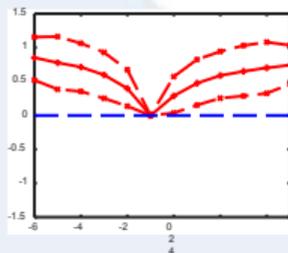
Monotone trend

No pre-trend

Multidimensional Variable omitida C_{it}



Proxy x_{it} si está disponible



Desempeno de diferentes estimadores

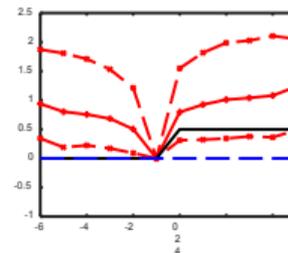
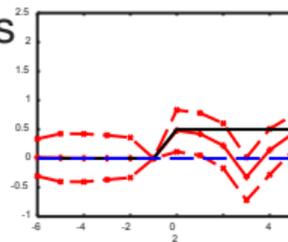
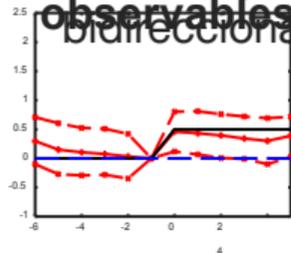
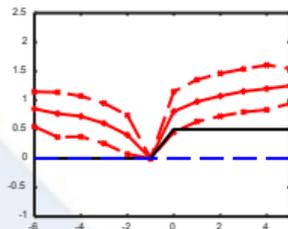
Mean-rev. trend
pre-trend

Monotone trend

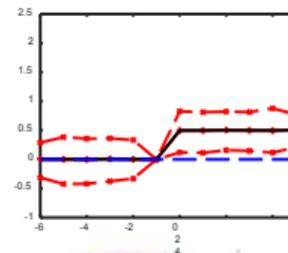
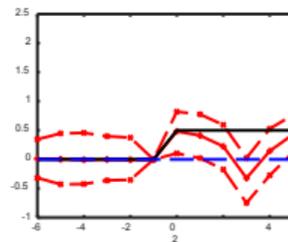
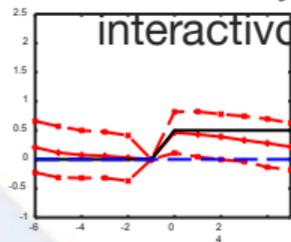
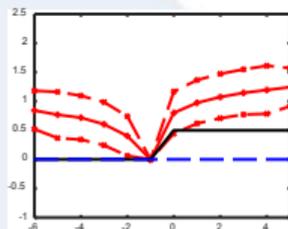
No

Multidimensiona

Condicional en
observables



Efectos fijos
interactivos



Desempenho de diferentes estimadores

Mean-rev. trend

Monotone trend

No

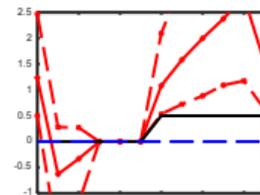
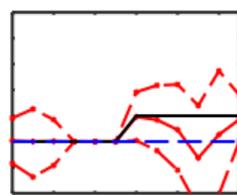
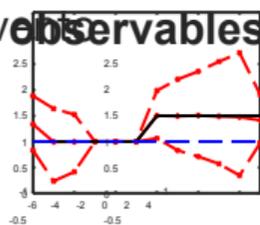
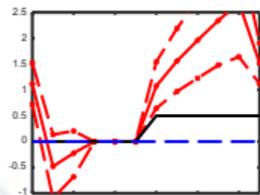
Multidimensiona

pre-trend

|

Exposições em tempo de

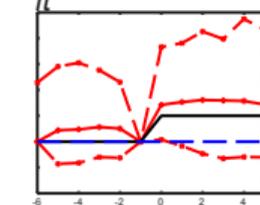
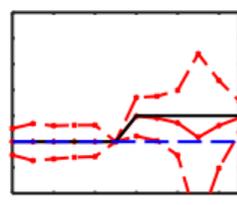
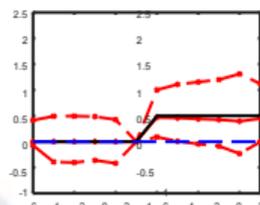
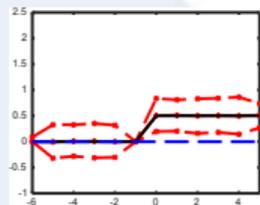
observáveis



Proxies e

Instrumentos

Instrumentando for x_{it} com adelantos de z_{it}



Conclusión

- ▶ Contribuciones de la presentación:
 1. Sugerimos mejoras a los gráficos de estudios de evento
 2. Discutimos algunos enfoques de identificación en presencia de variables omitidas
 3. Examinamos el desempeño de diferentes estimadores que corresponden a diferentes supuestos de identificación para varios procesos generadores de datos
- ▶ Ningún estimador tiene un buen desempeño para todos los procesos generadores de datos
- ▶ El desempeño de un estimador típicamente no puede evaluarse usando los datos disponibles
- ▶ Es importante motivar los supuestos del modelo usando economía



Escanee el

QR para versiones actualizadas del artículo y el paquete de Stata

Conferencias Stata LATAM 2022

Herramientas y aplicaciones estadísticas para Ciencia de Datos

Organiza



Conozca más sobre STATA
escaneando el código QR.

