Banco de México

Working Papers

N° 2024-09

Modelos de estudio de evento en panel: Una introducción al paquete xtevent que permite su estimación y visualización

Simon Freyaldenhoven
Federal Reserve Bank of Philadelphia

Christian B. Hansen
University of Chicago

Jorge Pérez Pérez
Banco de México

Jesse M. Shapiro
Harvard University and NBER

Constantino Carreto
Banco de México

Agosto de 2024

La serie de Documentos de Investigación del Banco de México divulga resultados preliminares de trabajos de investigación económica realizados en el Banco de México con la finalidad de propiciar el intercambio y debate de ideas. El contenido de los Documentos de Investigación, así como las conclusiones que de ellos se derivan, son responsabilidad exclusiva de los autores y no reflejan necesariamente las del Banco de México.

The Working Papers series of Banco de México disseminates preliminary results of economic research conducted at Banco de México in order to promote the exchange and debate of ideas. The views and conclusions presented in the Working Papers are exclusively the responsibility of the authors and do not necessarily reflect those of Banco de México.

Documento de Investigación
2024-09

Working Paper
2024-09

Modelos de estudio de evento en panel: Una introducción al paquete xtevent que permite su estimación y visualización[*]

Simon Freyaldenhoven[†]
Federal Reserve Bank of Philadelphia

Christian B. Hansen[‡]
University of Chicago

Jorge Pérez Pérez[§]
Banco de México

Jesse M. Shapiro[**]
Harvard University and NBER

Constantino Carreto[±]
Banco de México

Resumen: Los modelos de panel lineal y los "gráficos de estudio de evento" que los acompañan son herramientas populares para estudiar los efectos de políticas. En este artículo presentamos el paquete "xtevent" para Stata, que permite construir gráficos de estudio de evento siguiendo las sugerencias de Freyaldenhoven et al. (por publicarse). El paquete implementa varios procedimientos para estimar los efectos de políticas que están detrás de los gráficos, y permite el uso de variables de política no binarias y estimación ajustando por tendencias previas.
Palabras Clave: xtevent, modelos lineales de panel de datos, regresión de efectos fijos bidireccionales, tendencias previas, estudio de evento
JEL Classification: C23, C87, C53

Abstract: Linear panel models and the "event-study plots" that often accompany them are popular tools for learning about policy effects. In this paper, we introduce the "xtevent" package for Stata, which enables the construction of event-study plots following the suggestions in Freyaldenhoven et al. (forthcoming). The package implements various procedures to estimate the policy effects that underlie the plots, and allows for non-binary policy variables and estimation adjusting for pre-event trends.
Keywords: Xtevent, linear panel data models, two-way fixed effects regression, pre-trends, event study

*Agradecemos a nuestros dedicados asistentes de investigación por sus contribuciones. También agradecemos a los múltiples usuarios de xtevent por su retroalimentación. Agradecemos soporte financiero del programa SUPER RA en Harvard (Shapiro). Las opiniones expresadas en este documento son de los autores y no reflejan necesariamente las opiniones de la Reserva Federal de Philadelphia, la Reserva Federal, o las fuentes de financiamiento. Este artículo está en revisión en el Stata Journal.

† . Correo: simon.freyaldenhoven@phil.frb.org
‡ Booth School of Business. Correo: chansen1@chicagobooth.edu
§ Dirección General de Investigación Económica. Correo: jorgepp@banxico.org.mx
** Department of Economics. Correo: jesse_shapiro@fas.harvard.edu
± Dirección General de Investigación Económica. Correo: constantino.carreto@banxico.org.mx

1. Introducción

Los diseños de estudios de eventos en panel son comúnmente utilizados para analizar los efectos dinámicos de una política sobre una variable de resultado. Por ejemplo, Roth (2022) revisa las publicaciones en tres de las principales revistas económicas entre 2014 y 2018 y encuentra 70 artículos que incluyen diseños de estudios de eventos. Por su parte, Schmidheiny y Siegloch (2023) revisan 5 de las principales revistas de economía desde 1990 y observan un aumento gradual de la proporción de artículos que utilizan diseños de estudio de eventos, notando que su popularidad ha aumentado de manera considerable desde 2010, dado que esta proporción se cuadruplicó aproximadamente entre 2010 y 2018. Este uso creciente de diseños de estudios de eventos ha tenido lugar en una amplia gama de contextos: por ejemplo, para estudiar las respuestas de movilidad a las reformas fiscales (Martínez, 2022), el efecto del aumento del salario mínimo en la pobreza (Burkhauser et al., 2023), el efecto del cierre de escuelas en la salud mental de los padres (Gupta et al., 2024), el efecto de los conflictos armados en el comercio (Korovkin y Makarin, 2023), el impacto económico de la destitución de ayuntamientos (Fenizia y Saggio, 2024), y el efecto de la introducción de Internet rápido en el empleo (Hjort y Poulsen, 2019). La popularidad de este método plantea la necesidad de contar con herramientas computacionales que faciliten la implementación de los diseños de estudios de eventos.

Este artículo presenta `xtevent`, un paquete de Stata que permite la estimación de modelos lineales de panel con efectos dinámicos de política bajo diversos supuestos de identificación. Además, permite construir los gráficos de estudio de eventos correspondientes siguiendo las recomendaciones de Freyaldenhoven et al. (Forthcoming).

El interés principal de este estudio es conocer el efecto dinámico de una política escalar z_{it} sobre algún resultado y_{it} en un panel observacional de unidades $i \in \{1, ..., N\}$ observado en una secuencia de períodos $t \in \{1, ..., T\}$.Se considera el siguiente modelo:

$$y_{it} = \alpha_i + \gamma_t + q'_{it}\psi + \sum_{m=-G}^{M} \beta_m z_{i,t-m} + C_{it} + \varepsilon_{it}. \tag{1}$$

Aquí, α_i denota un efecto fijo unitario, γ_t un efecto fijo temporal, y q_{it} un vector de controles con coeficientes conformables ψ.

El escalar C_{it} denota una variable de confusión (potencialmente no observado) que puede estar correlacionado con la política, y el escalar ε_{it} representa un choque de política no observado. Los parámetros $\{\beta_m\}_{m=-G}^{M}$ encapsulan los efectos dinámicos de la

política. En concreto, el resultado en el momento t puede verse directamente afectado por el valor de la variable de política como máximo en $M \geq 0$ períodos antes de t y como máximo en $G \geq 0$ períodos después de t.

Los gráficos de estudio de eventos típicos que se utilizan para visualizar los efectos dinámicos de la política se basan en la siguiente variación de (1) (véase Freyaldenhoven et al. Forthcoming:

$$y_{it} = \sum_{k=-G-L_G}^{M+L_M-1} \delta_k \Delta z_{i,t-k} + \delta_{M+L_M} z_{i,t-M-L_M} + \delta_{-G-L_G-1}(1 - z_{i,t+G+L_G}) + \alpha_i + \gamma_t + q'_{it}\psi + C_{it} + \varepsilon_{it}, \tag{2}$$

donde Δ denota el operador de primeras diferencias. En (2), los parámetros $\{\delta_k\}_{k=-G-L_G-1}^{k=M+L_M}$ miden el efecto acumulativo de la política en diferentes horizontes (Schmidheiny y Siegloch, 2023).

El gráfico correspondiente del estudio de eventos muestra las estimaciones de los efectos acumulados del tratamiento en diferentes horizontes k. Así, el eje x corresponde a diferentes valores de k y el eje y y corresponde a las estimaciones de los efectos de la política $\left\{\widehat{\delta_k}\right\}_{k=-G-L_G-1}^{k=M+L_M}$. Nos referimos a k como el tiempo del evento (event-time), al vector $\delta = (\delta_{-G-L_G-1}, \ldots, \delta_{M+L_M})'$ como la trayectoria en el tiempo del evento (event-time path) del resultado, y a su contraparte estimada $\widehat{\delta}$ como la trayectoria estimada en el tiempo del evento (estimated event-time path).

Para permitir la visualización de sobreidentificación de información, la ecuación (2) incluye las estimaciones de los efectos acumulativos de la política en horizontes fuera del rango de horizontes en los que se considera que la política afecta el resultado Por ejemplo, es habitual descartar los efectos de la política en el momento t sobre el resultado en períodos anteriores a t ($G = 0$). Al incluir L_G períodos adicionales en (2), es posible visualizar las tendencias previas al evento ("pre-trends") que generalmente son inconsistentes con el modelo en (1) (Freyaldenhoven et al., 2019). Del mismo modo, la ecuación de estimación en (2) permite visualizar el efecto acumulativo estimado para L_M periodos adicionales después de que se da por sentado que el efecto acumulativo del tratamiento es constante en (1).

Aunque (2) permite variables de política generales (por ejemplo, no binarias) z_{it}, considerar el caso particular de *adopción escalonada*, es esclarecedor. La adopción escalonada quiere decir que la política es binaria, todas las unidades comienzan sin la

política; y una vez que una unidad dada adopta la política nunca se revierte. Entonces, $\Delta z_{i,t-k}$ es un indicador de si la unidad i adoptó la política exactamente k periodos antes del periodo t, $z_{i,t-M-L_M}$ es un indicador de si la unidad i adoptó al menos $M + L_M$ periodos antes del periodo t, y $(1 - z_{i,t+G+L_G})$ es un indicador de si la unidad i adoptará más de $G + L_G$ periodos después del periodo t.

El paquete complementa muchas otras contribuciones recientes a la estimación y visualización de estudios de eventos de panel en Stata, como `csdid` (Rios-Avila et al., 2021), `did_imputation` (Borusyak, 2023), `didmultiplegt` (de Chaisemartin et al., 2019), `did2s` (Butts, 2023), `eventdd` (Clarke y Tapia-Schythe, 2021), `eventstudyinteract` (Sun, 2023), `jwdid` (Rios-Avila, 2022), `lpdid` (Busch y Girardi, 2023), `staggered` (Cáceres Bravo, 2023), y `wooldid` (Hegland, 2023),), así como el comando nativo `xthdidregress` introducido recientemente en Stata 18. Muchos de ellos se centran en el caso de la adopción escalonada y, por defecto, requieren que el usuario precise un período de tratamiento específico de la unidad, el tiempo relativo al tratamiento o un indicador para las observaciones. sujetas a tratamiento. Por el contrario, el paquete utilizado permite variables de política general z_{it}, como variables continuas, los que hace posible hacer estimaciones y visualizar en una amplia gama de escenarios fuera de la adopción escalonada. Se observa que el paquete de este estudio puede utilizarse en el caso de adopción escalonada estableciendo z_{it} igual a un indicador específico de período unitario para los períodos posteriores al tratamiento. El paquete también permite la estimación con tendencias previas al evento utilizando enfoques basados en la extrapolación de tendencias (Dobkin et al., 2018) o variables proxy (Freyaldenhoven et al., 2019). Además, incluye herramientas para mejorar los gráficos de estudio de eventos y facilitar su interpretación, como el cálculo de bandas de confianza uniformes y el trazado de trayectorias de confusión plausibles en línea con la trayectoria de tiempo de evento estimada. Existen ventajas en las implementaciones diseñadas para garantizar las propiedades económicas deseables en entornos más especializados, como la adopción escalonada. El paquete incorpora uno de estos procedimientos como opción si la política sigue efectivamente la adopción escalonada.

En la siguiente sección se analizan brevemente varias estrategias de estimación para (2), proporcionando más detalles sobre la construcción de las gráficas de estudio de evento correspondientes y presentando algunas características adicionales del paquete `xtevent` . La sección 3 ilustra el uso del paquete en datos simulados de Freyaldenhoven et al. (Forthcoming) (próxima publicación) y estima el efecto de una reforma fiscal utilizando datos de Martínez (2022). Por último, el apéndice ofrece una descripción más

detallada de la sintaxis y las opciones del paquete `xtevent`. El apéndice también incluye detalles adicionales sobre la implementación y funcionalidad del paquete.

2. Métodos

2.1. Estrategias de estimación

En general, la identificación de los parámetros δ requerirá algún tipo de restricción sobre la forma en que las variables observables y latentes se relacionan con la variable de confusión C_{it} y la política z_{it}. La restricción adecuada dependerá del entorno económico y normalmente no puede aprenderse a partir de los datos. A su vez, la elección de la restricción determinará qué tipo de estimador es apropiado para estimar δ (véase Freyaldenhoven et al. Forthcoming para un análisis más detallado). El paquete `xtevent` incluye los siguientes estimadores:

Estimador de efectos fijos bidireccionales. Si $C_{it} = 0$, la ecuación (2) puede estimarse por mínimos cuadrados ordinarios (MCO) utilizando un estimador estándar de efectos fijos bidireccionales. Con un solo grupo de efectos fijos, `xtevent` utiliza `areg` para la estimación. `xtevent` permite además la estimación utilizando `xtreg` o con el comando `reghdfe` (Guimarães y Portugal, 2010; Correia, 2016, 2019) para permitir efectos fijos múltiples o de alta dimensión.

Control de las tendencias específicas de unidad. Si $C_{it} = \lambda_i' f(t)$, donde $f(\cdot)$ es un conjunto conocido de funciones base de baja dimensión (por ejemplo, $f(t) = t$, entonces (2) puede estimarse incluyendo tendencias temporales específicas a cada unidad. Estas pueden incluirse en la regresión utilizando variables factoriales (por ejemplo, `i.crosssectionid#c.time`)) o absorbiéndolas mediante `reghdfe`.

Control de las tendencias por tiempo de evento. Si C_{it} puede escribirse como

$$C_{it} = \tilde{\alpha}_i + \tilde{\gamma}_t + q_{it}'\tilde{\psi} + \sum_m \phi' f(m) z_{i,t-m} \tag{3}$$

para un conjunto conocido de funciones de base $f(\cdot)$ y parámetros desconocidos $\tilde{\alpha}_i, \tilde{\gamma}_t$ y $\tilde{\psi}$, entonces la ecuación (2) puede estimarse incluyendo los términos apropiados de (3) directamente en un modelo de regresión, o por método de los momentos generalizados (GMM, por sus siglas en inglés) en un segundo paso tras la estimación de (2) mediante efectos fijos bidireccionales.

Intuitivamente, supóngase que una tendencia de tiempo de evento puede aproximar la variable de confusión. En ese caso, se puede conocer la tendencia en periodos en los

que la política está inactiva y extrapolarla a periodos posteriores. Las diferencias entre la variable de resultado y la tendencia extrapolada son entonces informativas de los efectos de la política (Dobkin et al., 2018). Por ejemplo, considérese un escenario de adopción escalonada en el que la variable de confusión sigue una tendencia lineal en el tiempo; esta tendencia comienza tres periodos antes de que se active la política y continúa durante tres periodos después. Esta situación se puede representar tomando $f(m) = 1$ si $m \in [-3, 3]$ y $f(m) = 0$ en caso contrario. En este caso, se puede extrapolar la tendencia a los periodos posteriores a la adopción y restarla para tener en cuenta la variable de confusión. Se permite que z_{it} sea continua y que la adopción no sea escalonada. Si se cumple la ecuación (3) entonces el estimando de un estimador estándar de efectos fijos bidireccionales de la ecuación (2) está dado por

$$d_k = \begin{cases} \phi' f_k, & \text{if } k < -G \\ \delta_k + \phi' f_k, & \text{if } -G \leq k \leq M \\ \delta_M + \phi' f_k, & \text{otherwise,} \end{cases} \tag{4}$$

donde $f_k = \sum_{m=-\infty}^{k} f(m)$.

Dadas las estimaciones de d_k, $\widehat{d_k}$, se puede recuperar los parámetros de tendencia ϕ a partir de las $\widehat{d_k}$ en los L_G periodos no afectados. Sea $T_G \leq L_G$ el número de períodos anteriores a G utilizados para estimar los parámetros de la tendencia y $T_M \leq M$ el número de períodos "posteriores al evento" en los que la tendencia está activa. Se da por sentado que $f_k \neq 0$ para $k \in [-G - T_G, T_M]$ y 0 en caso contrario. Podemos recuperar los parámetros de tendencia utilizando T_G condiciones de momentos $\widehat{d_k} - \phi' f_k = 0$ para $k = -G - T_G, \ldots, -G - 1$. A continuación, es posible calcular una trayectoria de tiempo de evento estimada ajustada $\widehat{\delta}$ que tenga en cuenta la variable de confusión restando $\phi' f_k$ de los coeficientes no ajustados d_k para $k \in [-G - T_G, T_M]$. El Apéndice B ofrece más detalles sobre este estimador.

Estimación IV con múltiples variables "proxy". Si se dispone de múltiples variables adicionales que puedan servir como variables "proxy" para la variable de confusión C_{it}, , de forma que

$$x_{it} = \alpha_i^x + \gamma_t^x + \phi^x q_{it} + \Xi^x C_{it} + u_{it} \tag{5}$$

donde los parámetros $\alpha_i^x, \gamma_t^x, \Xi^x$ son desconocidos y u_{it} es un vector no observado (el cual no está correlacionado entre las variables "proxy"), entonces `xtevent` permite estimar

la ecuación (2) por mínimos cuadrados en dos etapas, incluyendo una de los proxies en (2), y utilizando la otra variable proxy como instrumento excluido (Freyaldenhoven et al., 2019).

Estimación IV con una única variable proxy. Si solo se dispone de una única variable adicional que pueda servir como proxy para la variable de confusión C_{it}, y si el error en la ecuación (5) es condicionalmente independiente de la media de la política, entonces `xtevent` permite estimar la ecuación (2) utilizando un estimador de mínimos cuadrados en dos etapas, instrumentando la proxy con adelantos de la variable de política (Freyaldenhoven et al., 2019).

En principio, todos los adelantos de la variable de política más alejados que G son potenciales instrumentos. Para seleccionar una opción por defecto, `xtevent` estima (2) mediante efectos fijos bidireccionales, con la proxy como variable de resultado. El instrumento excluido por defecto para la estimación de (2) es entonces la variable con el mayor valor absoluto del estadístico t entre los adelantos de la diferencia de la variable de política $\{\Delta z_{i,t+k}\}_{k=G+1}^{G+L_G}$ y $z_{i,t+G+L_G}$.

2.2. Gráficos de estudio de eventos

El comando auxiliar `xteventplot` incluye la funcionalidad para visualizar los gráficos de estudio de eventos basados en cualquiera de los estimadores de la sección anterior. Además, incluye las mejoras de estos gráficos sugeridas en Freyaldenhoven et al. (Forthcoming).

Normalización. Debido a que las variables políticas en (2) son colineales, se requiere una normalización para identificar la trayectoria tiempo de evento $\{\delta_k\}_{k=-G-L_G-1}^{k=M+L_M}$ (Schmidheiny y Siegloch, 2023; Freyaldenhoven et al., Forthcoming). `xtevent` normaliza $\delta_{-1} = 0$ por defecto. En el caso de la adopción escalonada, esta normalización implica que los coeficientes trazados pueden interpretarse como efectos estimados relativos al periodo anterior a la promulgación de la política.

Nivel de la variable de resultado. Para facilitar la interpretación de los efectos políticos estimados `xtevent` incluye una etiqueta entre paréntesis para el coeficiente normalizado que refleja la media de la variable dependiente. Por ejemplo, en el caso de la adopción escalonada, según la normalización por defecto, la etiqueta corresponde a la media muestral de y_{it} un periodo antes de la adopción. En términos más generales, la

etiqueta corresponde al valor

$$\frac{\sum_{(i,t):\Delta z_{i,t-k^\star}\neq 0} y_{it}}{|(i,t):\Delta z_{i,t-k^\star}\neq 0|}$$

donde $k^\star$ corresponde al coeficiente normalizado de tiempo de evento $\delta_{k^\star}$.

Inferencia uniforme. Además de los intervalos de confianza puntuales estándar para los coeficientes δ_k, `xtevent` permite trazar bandas de confianza uniformes sup-t (Freyberger y Rai, 2018; Montiel Olea y Plagborg-Møller, 2019). La inclusión de estas bandas permite realizar pruebas visuales de hipótesis sobre toda la trayectoria de coeficientes en lugar de coeficientes únicos solamente.

Sobre-identificación y pruebas. La ecuación de estimación en (2) incluye L_G períodos adicionales antes de la adopción de la política para visualizar posibles tendencias previas. En la práctica, la evidencia de tales tendencias previas suele considerarse como indicio de la presencia de un factor de confusión que invalida el diseño de investigación. Asimismo, la ecuación de estimación en (2) también incluye L_M periodos adicionales después de la finalización de los efectos de la política para evaluar si los efectos dinámicos se han nivelado después de los M periodos postulados para los que la política tiene un efecto directo. `xtevent` muestra los valores p de las pruebas de Wald para "tendencias previas" ($\delta_k = 0$ para $-G - L_G - 1 \leq k < -G$) y para "nivelación" de efectos dinámicos ($\delta_M = \delta_{M+k}$ para $0 < k \leq L_M$). El comando auxiliar `xteventtest` permite probar hipótesis adicionales, como las hipótesis sobre efectos acumulativos, si los efectos son constantes y si siguen tendencias lineales.

Trayectoria más suave de las variables de confusión consistente con las estimaciones.

Para ayudar a visualizar si una variable de confusión puede explicar de forma plausible toda la dinámica temporal de la variable de resultado, `xtevent` permite añadir una representación de la trayectoria "más suave" de las variables de confusión coincidente con la ausencia de efectos políticos. La variable de confusión seleccionada"más suave" será el polinomio menos "ondulante" el tiempo de evento que pasa por la región de confianza de Wald de la trayectoria del tiempo de evento. La idea es que si existe una trayectoria "suave", esto sugiere que una variable de confusión podría explicar plausiblemente toda la trayectoria de tiempo de evento del resultado, incluso en ausencia de efectos de política. Por otra parte, si no existe una trayectoria "suave" esto puede indicar que una variable de confusión no puede explicar plausiblemente toda la trayectoria temporal del resultado y, por lo tanto, si la política afecta al resultado. En el Apéndice D se describe detalladamente el cálculo de la trayectoria más suave.

Gráficos superpuestos. `xteventplot` permite gráficos de estudio de eventos con superposiciones en diferentes escenarios de estimación. Para la estimación con tendencias de tiempo de evento, `xteventplot` crea gráficos superponiendo la tendencia. En el caso de la estimación por variables instrumentales (IV) con proxies `xteventplot` permite superponer la dinámica implícita en la variable proxy. Además, `xteventplot` permite superponer un modelo de efectos constantes para evaluar si los efectos de la política se mantienen constantes a lo largo del tiempo.

2.3. Características adicionales

El paquete incluye las siguientes funciones adicionales.

Imputación de los valores faltantes en la variable de política y sus adelantos y rezagos. Dado que la ecuación (1) incluye adelantos y rezagos de la variable de política z_{it} y su primera diferencia, la muestra de estimación puede ser menor a la muestra completa disponible (Schmidheiny y Siegloch, 2023). En general, si la variable de resultado se observa para $t \in \{\underline{t}, \ldots, \overline{t}\}$, es necesario observar la variable de política z_{it} desde $\underline{t} - G - L_G$ hasta $\overline{t} + M + L_M - 1$ para evitar que se pierdan observaciones de la muestra de estimación. En una configuración típica, esto puede implicar que se tiene que restringir la ventana de estimación para calcular los adelantos y rezagos necesarios de z_{it}. Sin embargo, el usuario puede tener información adicional que permita la imputación de la variable de política.

`xtevent` permite los siguientes esquemas de imputación:

1. Si se declara que la variable de política sigue la adopción escalonada, `xtevent` puede:

   a) Imputar automáticamente cualquier valor que falte en la variable de política *fuera* del intervalo de datos observados, suponiendo que no hay cambios de política fuera del período de la muestra. Por ejemplo, si se observa $z_{jt} = 1$, en caso de una adopción escalonada, esto implica que $z_{js} = 1$ para $s > t$. En la sección 3.1 se ofrece un ejemplo de esta función.

   b) Imputar automáticamente los valores faltantes de la variable de política *dentro del rango de datos* observado. Por ejemplo, si $z_{jt} = 1$ y $z_{j,t+2} = 1$, ubajo adopción escalonada, esto implica que $z_{j,t+1} = 1$. Ejemplos de esta función se proporcionan en el Apéndice E.

2. Incluso en ausencia de adopción escalonada, si se declara que la variable de política no cambia de valor fuera de la muestra observada, `xtevent` puede imputar automáticamente z_{it} fuera de la muestra observada. Por ejemplo, si la muestra comienza en $\underline{t}$ y $z_{1\underline{t}} = 0$ para la unidad 1, se establece $z_{it} = 0$ para $t \in \{\underline{t} - G - L_G, ..., \underline{t} - 1\}$.

Estimación con datos de corte transversales repetidos. Algunas configuraciones implican datos de corte transversales repetidos en lugar de datos de panel, con un tratamiento que varía a un nivel superior de agregación. Por ejemplo, se puede tener una serie de cortes transversales repetidos de individuos para cada estado de EE.UU s y una política a nivel estatal z_{st}. En este entorno, la estimación de (2) con efectos fijos individuales es inviable, pero `xtevent` permite la estimación de un modelo relacionado con efectos fijos por estado:

$$y_{it} = \sum_{k=-G-l_G}^{M+L_M-1} \delta_k \Delta z_{s(i),t-k} + \delta_{M+L_M} z_{s(i),t-M-L_M} + \delta_{-G-L_G-1}(1 - z_{s(i),t+G+L_G}) + \alpha_{s(i)} + \gamma_t + q'_{it}\psi + C_{s(i)t} + \varepsilon_{it}, \tag{6}$$

donde los parámetros $\alpha_{s(i)}$ son efectos fijos correspondientes al estado s al que pertenece la unidad i

Una alternativa (Amemiya, 1978; Hansen, 2007) es estimar una regresión de y_{it} sobre un conjunto de indicadores de estado-tiempo, más cualquier variable de control q_{it} que varíe a nivel individual, y luego estimar (2) con los efectos estimados de estado-tiempo como variables dependientes, e incluyendo efectos fijos de estado, efectos fijos de tiempo y controles que varíen a nivel de estado. El comando auxiliar `get_unit_time_effects` facilita este enfoque. El Apéndice F proporciona un ejemplo de este comando .

Efectos heterogéneos del tratamiento en contextos de adopción escalonada. El modelo de la ecuación (1) supone que el efecto causal de la política es homogéneo entre unidades i. La literatura reciente ha destacado que si los efectos del tratamiento son heterogéneos según el tiempo de tratamiento, entonces los efectos estimados con la ecuación (1) pueden no ser promedios adecuadamente ponderados de los efectos del tratamiento a nivel de cohorte (Athey y Imbens, 2022; Callaway y Sant'Anna, 2021; Goodman-Bacon, 2021; Sun y Abraham, 2021). Sun y Abraham (2021) proponen estimar estudios de eventos para cada cohorte tratada por separado, comparando cada una con una cohorte no tratada, y luego promediar los efectos, ponderando por el porcentaje de unidades tratadas en cada cohorte, para llegar a un efecto del tratamiento sobre

las unidades tratadas. Para el caso de efectos fijos bidireccionales, y bajo adopción escalonada, `xtevent` permite la estimación de efectos específicos de cada cohorte, en cuyo caso informa de una media ponderada por el número de observaciones tratadas en cada cohorte, que es una estimación de un efecto de tratamiento medio ponderado de las observaciones tratadas bajo los supuestos discutidos en Sun y Abraham (2021). La sección 3 ofrece un ejemplo de estimación con efectos de tratamiento heterogéneos y Apéndice A da detalles sobre la estimación.

3. Ejemplos

Esta sección ofrece dos ejemplos de uso de `xtevent` En primer lugar, se muestra la funcionalidad básica del paquete utilizando datos simulados de Freyaldenhoven et al. (Forthcoming). Después se muestran opciones adicionales utilizando datos reales de Martínez (2022).

3.1. Ejemplo de datos simulados

Primero se utilizan los datos de "Jump at the time of the event" de Freyaldenhoven et al. (2022). Los datos son un panel equilibrado de 2,000 observaciones de 50 unidades observadas a lo largo de 40 periodos, en los que la política afecta inicialmente a las unidades en periodos diferentes. Las unidades son tratadas aleatoriamente en el periodo de muestra. El coeficiente de la variable de tratamiento es uno. Para empezar se debe cargar la base de datos especificando las variables de unidad y tiempo.

```
. use simulation_data_dynamic.dta, clear

. xtset id t

Panel variable: id (strongly balanced)
 Time variable: t, 1 to 40
         Delta: 1 unit
```

El cuadro que aparece abajo es un vistazo de la base de datos. La variable `id` indexa las unidades transversales, `t` indexa el tiempo, `y` es la variable de resultado, `z` es la variable de política y `x` es una variable de control.

```
. list id t z y x if id==2 & t<=10, noobs
```

id	t	z	y	x
2	1	0	42.02239	-.0102958

2	2	0	42.83109	.1713373
2	3	0	42.82665	-.197851
2	4	1	42.59289	-.5887834
2	5	1	43.3557	-.3722385
2	6	1	42.74355	.355894
2	7	1	42.82405	-2.047098
2	8	1	42.34953	-.3757658
2	9	1	42.14841	-1.976451
2	10	1	41.79388	-1.16444

Nótese que la variable `t` es un momento en el calendario, no un evento. `xtevent` no requiere normalizar la variable temporal, ya que la especificación de la ecuación (2) permite variables políticas discretas o continuas y cambios únicos o múltiples.

Funcionalidad básica

La ecuación (2) se estima estableciendo $M + L_M = 5$ y $G + L_G = 5$, por lo que analizamos los efectos de la política en el resultado cinco períodos antes y cinco períodos después de la adopción de la política.

Para estimar un estudio de eventos en panel básico con efectos dinámicos de la variable de política `z` sobre `y` utilizando `x` como variable de control y graficar los resultados, escribimos:

```
. xtevent y x, panelvar(id) timevar(t) policyvar(z) window(5) impute(nuchange) ///
> plot

No proxy or instruments provided. Implementing OLS estimator

Linear regression, absorbing indicators             Number of obs     =  2,000
Absorbed variable: id                               No. of categories =     50
                                                    F(52, 1898)       =   2.81
                                                    Prob > F          = 0.0000
                                                    R-squared         = 0.0980
                                                    Adj R-squared     = 0.0500
                                                    Root MSE          = 0.7181
```

y	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
_k_eq_m6	.2008833	.1138905	1.76	0.078	-.0224803	.424247
_k_eq_m5	.2935611	.1499538	1.96	0.050	-.0005306	.5876528
_k_eq_m4	.1172962	.1490125	0.79	0.431	-.1749493	.4095416
_k_eq_m3	.0874992	.1472188	0.59	0.552	-.2012285	.3762268
_k_eq_m2	.0415972	.1472521	0.28	0.778	-.2471958	.3303902
_k_eq_p0	.8160009	.1472673	5.54	0.000	.5271782	1.104824
_k_eq_p1	.8400974	.1475249	5.69	0.000	.5507695	1.129425
_k_eq_p2	.5699733	.1497108	3.81	0.000	.2763583	.8635883
_k_eq_p3	.2665183	.1507838	1.77	0.077	-.0292011	.5622377
_k_eq_p4	.0915933	.1547402	0.59	0.554	-.2118855	.3950721

```
      _k_eq_p5 |    .091785   .1564831     0.59   0.558     -.215112     .398682
      _k_eq_p6 |   .1405872   .1192801     1.18   0.239    -.0933466     .374521
             x |   .0076964   .0302145     0.25   0.799    -.0515607    .0669535
               |
             t |
(output omitted)
               |
         _cons |   41.91144   .1507874   277.95   0.000     41.61572    42.20717
------------------------------------------------------------------------------
F test of absorbed indicators: F(49, 1898) = 1.082             Prob > F = 0.325
```

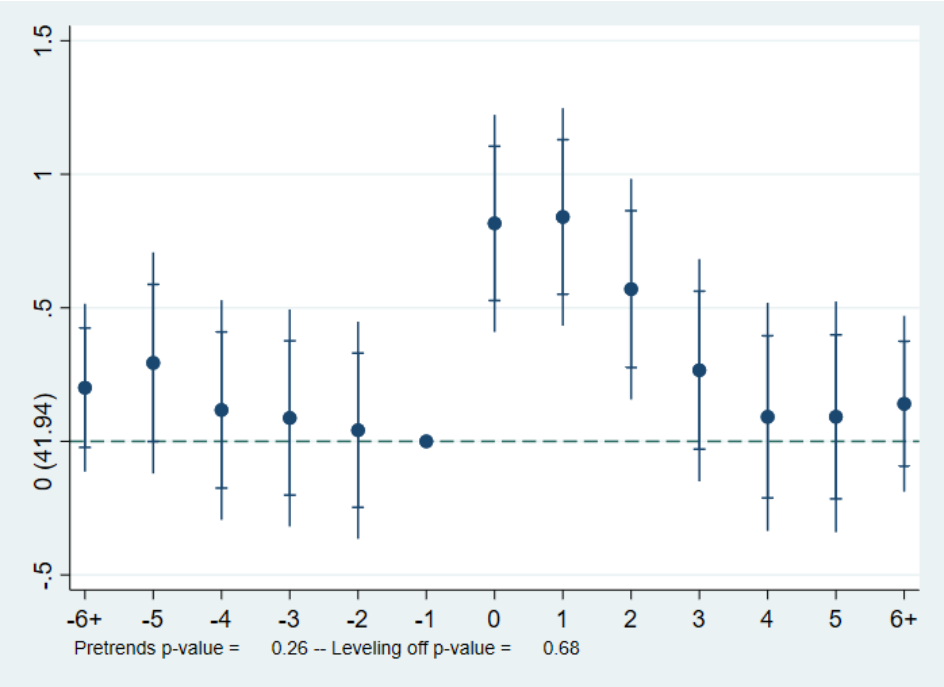
Las opciones `panelvar` y `timevar` indican las dimensiones transversal y temporal de la base de datos. La opción `window` específica los periodos de tiempo de evento que habrá que incluir. Al especificar la opción `impute(nuchange)`, se le pide a `xtevent` que asuma que la política no cambia fuera de la ventana de estimación y que impute las anticipaciones (leads) y rezagos (lags) de la variable de política en consecuencia. Discutimos esquemas alternativos de imputación en el Apéndice E

`xtevent` crea automáticamente variables indicadoras para tiempos de evento entre -5 a 5, denotadas como `_k_eq_m5`,`_k_eq_m4`,...,`_k_eq_p5`, además de variables indicadoras de los puntos extremos para tiempos de evento ≤ -6 y ≥ 6 (`_k_eq_m6`,`_k_eq_p6`) e incluye estas variables como independientes en la regresión. El coeficiente normalizado para el tiempo de evento $k = -1$ se omite. Por defecto, `xtevent` incluye efectos fijos de unidad y de tiempo en la regresión, aunque estos pueden excluirse mediante las opciones `note` y `nofe`. En este caso, `xtevent` utiliza `areg` para estimar la regresión, por lo que los efectos fijos de unidad no se reportan.

La salida predeterminada incluye errores estándar convencionales.

La opción `plot` solicita un gráfico de estudio de eventos después de la estimación, como se muestra en la figura 1. El gráfico también puede obtenerse escribiendo `xteventplot` después de la llamada a `xtevent` . El gráfico por defecto muestra los valores de los coeficientes estimados junto con intervalos de confianza puntuales (bigotes internos) y bandas de confianza sup-t para toda la trayectoria de tiempo de evento (picos exteriores). Por defecto, `xtevent` normaliza $\delta_{-1} = 0$, y el eje y incluye un valor en paréntesis que indica la media de la variable dependiente un periodo antes de la adopción. En la parte inferior, el gráfico incluye valores p para una prueba de tendencias previas y una prueba de nivelación. Con la opción `reghdfe`, el usuario puede pedir a `xtevent` que estime el modelo utilizando el comando `reghdfe` de Correia (2016). El usuario también puede especificar variables adicionales para ser absorbidas a través de la opción `addabsorb(varlist)`. Para pedir a `xtevent` que estime con `reghdfe` y absorba adicionalmente la variable `eta_r2`, hay que escribir:

Figura 1: Gráfico de estudio de eventos



Fuente: Elaborado por los autores con datos de Freyaldenhoven et al. (2022).

```
. gen eta_r2=round((eta_r+1)*2)

. qui xtevent y x, panelvar(id) timevar(t) policyvar(z) window(5) ///
> reghdfe addabsorb(eta_r2) impute(nuchange)
```

Elegir diferentes ventanas

También es posible especificar una ventana asimétrica. Por ejemplo, para una estimación de 4 periodos anteriores al evento, 7 periodos posteriores al evento y dos puntos finales:

```
. qui xtevent y x, panelvar(id) timevar(t) policyvar(z) window(-4 7) ///
>  impute(nuchange)
```

Con la notación de la ecuación (2), corresponde a establecer G, el número de periodos anteriores al evento en los que pueden producirse efectos anticipados, en 0. L_G en 4 permitiendo visualizar los efectos previos en los cuatro períodos anteriores al evento. De manera análoga, se establece M en 7, lo que indica que se analizan siete períodos posteriores al evento para capturar efectos retardados, y L_M en 1 agregando un período

adicional para evaluar si los efectos se estabilizan con el tiempo. En resumen, esta configuración es equivalente a especificar explícitamente los valores de G, L_G, M, y L_M:

```
. qui xtevent y x, panelvar(id) timevar(t) policyvar(z) pre(0) overidpre(4) ///
>  post(7) overidpost(1) impute(nuchange)
```

Ajuste de tendencia lineal

La opción `trend(#1 [, subopt])` permite extrapolar una tendencia lineal en el tiempo de evento desde el periodo #1 antes del cambio de política, como en Dobkin et al. (2018). El efecto estimado de la política es la desviación de la tendencia lineal extrapolada. La tendencia lineal de tiempo de evento se estima utilizando tres periodos previos al evento $(-3, -2, -1)$ y la subopción `method(gmm)` usando un estimador de método generalizado de momentos. Los datos superpuestos también se guardan para hacer su representación gráfica:

```
. xtevent y x, panelvar(id) timevar(t) policyvar(z) window(5) ///
>  impute(nuchange) trend(-3, method(gmm) saveoverlay)

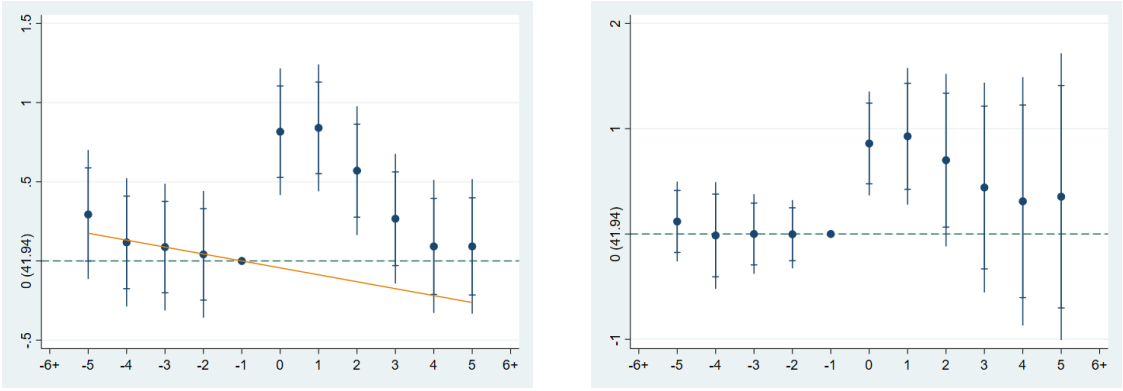
No proxy or instruments provided. Implementing OLS estimator

Linear regression, absorbing indicators            Number of obs     =  2,000
Absorbed variable: id                              No. of categories =     50
                                                   F(52, 1898)       =   2.81
                                                   Prob > F          = 0.0000
                                                   R-squared         = 0.0980
                                                   Adj R-squared     = 0.0500
                                                   Root MSE          = 0.7181
```

y	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
_k_eq_m6	-.0178682	.3123965	-0.06	0.954	-.6305449	.5948085
_k_eq_m5	.1185599	.1499569	0.79	0.429	-.1755379	.4126576
_k_eq_m4	-.0139547	.2002518	-0.07	0.944	-.4066915	.378782
_k_eq_m3	-1.45e-06	.1496187	-0.00	1.000	-.2934359	.293433
_k_eq_m2	-.0021531	.1281336	-0.02	0.987	-.2534507	.2491444
_k_eq_p0	.8597512	.1953144	4.40	0.000	.4766977	1.242805
_k_eq_p1	.927598	.256436	3.62	0.000	.4246719	1.430524
_k_eq_p2	.7012242	.3242324	2.16	0.031	.065335	1.337113
_k_eq_p3	.4415195	.3940669	1.12	0.263	-.3313303	1.214369
_k_eq_p4	.3103449	.4663886	0.67	0.506	-.6043433	1.225033
_k_eq_p5	.3542868	.5386234	0.66	0.511	-.7020694	1.410643
_k_eq_p6	.4468393	.6017334	0.74	0.458	-.733289	1.626968
x	.0076964	.0302145	0.25	0.799	-.0515607	.0669535
t						
2	.090792	.1439829	0.63	0.528	-.1915894	.3731734
(*output omitted*)						

Para visualizar las estimaciones originales, la tendencia extrapolada y las estimaciones ajustadas a la tendencia lineal, `xtevent` puede producir un gráfico superpuesto: se usa `xteventplot` para producir un gráfico superpuesto con la tendencia extrapolada y un gráfico ajustado. `xtevent` excluye los puntos finales en ambas figuras. La figura 2 muestra los gráficos superpuestos y ajustados del estudio de eventos..

Figura 2: Ajuste de la tendencia lineal



(a) Superposición de línea de extrapolación (b) Substracción de la tendencia extrapolada

Fuente: Elaborado por los autores con datos de Freyaldenhoven et al. (2022).

Ajuste de las tendencias previas con variables proxy

`xtevent` permite la estimación cuando hay una tendencia previa (pre-trend) presente, utilizando el estimador de variables instrumentales propuesto por Freyaldenhoven et al. (2019). Para ello, es necesario especificar la opción `proxy(varname)` para indicar la(s) variable(s) proxy del factor de confusión. Por defecto, `xtevent` xtevent realiza una regresión de la variable proxy sobre las variables principales de la variable de política diferenciada y elige la variable principal con el estadístico t absoluto más alto para utilizarla como instrumento para la variable proxy. Alternativamente, el usuario puede especificar una variable principal específica o diferentes variables a utilizar como instrumentos en la opción `proxyiv(string)`. `xtevent` utiliza `xtivregress` para estimar este modelo.

```
. xtevent y, panelvar(id) timevar(t) policyvar(z) window(5) impute(nuchange) ///
> proxy(x)

Proxy for the confound specified. Implementing FHS estimator
```

```
proxyiv=select. Selecting lead order of differenced policy variable to use as inst
> rument.

Lead 4 selected.

The coefficient at -1 is normalized to zero.

For estimation with proxy variables, an additional coefficient needs to be normali
> zed to zero.

The coefficient at -4 was selected to be normalized to zero.

Fixed-effects (within) IV regression           Number of obs     =      1,800
Group variable: id                             Number of groups  =         50

R-squared:                                     Obs per group:
     Within  =      .                                          min =         36
     Between = 0.0557                                         avg =       36.0
     Overall = 0.0335                                         max =         36

                                               Wald chi2(47)     =   5.83e+06
corr(u_i, Xb) = -0.0155                        Prob > chi2       =     0.0000
```

y	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
x	.3772094	.5069244	0.74	0.457	-.6163441	1.370763
_k_eq_m6	.1009198	.1046026	0.96	0.335	-.1040974	.3059371
_k_eq_m5	.2344818	.1356115	1.73	0.084	-.0313119	.5002755
_k_eq_m3	.0187142	.1328999	0.14	0.888	-.2417649	.2791932
_k_eq_m2	.0189579	.13511	0.14	0.888	-.2458528	.2837686
_k_eq_p0	.8122546	.1805533	4.50	0.000	.4583766	1.166133
_k_eq_p1	.9546866	.2887189	3.31	0.001	.3888079	1.520565
_k_eq_p2	.7687619	.3744836	2.05	0.040	.0347875	1.502736
_k_eq_p3	.4715529	.4518123	1.04	0.297	-.413983	1.357089
(output omitted)						
32	-.1106802	.1948427	-0.57	0.570	-.492565	.2712045
33	-.0372759	.2005944	-0.19	0.853	-.4304338	.3558819
34	-.0579759	.1700988	-0.34	0.733	-.3913634	.2754116
35	-.1346377	.1744911	-0.77	0.440	-.476634	.2073585
36	.2059306	.1766346	1.17	0.244	-.1402669	.5521281
_cons	41.96931	.1359876	308.63	0.000	41.70278	42.23584
sigma_u	.12123774					
sigma_e	.74115015					
rho	.02606123	(fraction of variance due to u_i)				

```
 F test that all u_i=0: F(49,1703) =     0.88                     Prob > F    = 0.7137

Endogenous: x
Exogenous:  _k_eq_m6 _k_eq_m5 _k_eq_m3 _k_eq_m2 _k_eq_p0 _k_eq_p1 _k_eq_p2
            _k_eq_p3 _k_eq_p4 _k_eq_p5 _k_eq_p6 2.t 3.t 4.t 5.t 6.t 7.t 8.t
            9.t 10.t 11.t 12.t 13.t 14.t 15.t 16.t 17.t 18.t 19.t 20.t 21.t
            22.t 23.t 24.t 25.t 26.t 27.t 28.t 29.t 30.t 31.t 32.t 33.t 34.t
            35.t 36.t _fd4__00000M
```

`xteventplot` puede crear varios gráficos para ilustrar este estimador. Primero se utiliza `xteventplot, y` para crear un gráfico de estudio de evento no ajustado para el

resultado. Este gráfico se muestra en el panel a) la figura 3. En segundo lugar, usando la opción `proxy`, se ilustra la dinámica de la variable proxy (que se muestra en el panel b) de la figura 3). En tercer lugar, utilizando la opción `overlay(iv)`, se crea un gráfico que alinea la dinámica del proxy y el resultado entre el coeficiente utilizado como instrumento y el coeficiente normalizado (como muestra en el panel c) de la figura 3). Por último, utilizando `xteventplot` y sin opciones adicionales, se crea un gráfico que muestra los coeficientes del resultado después de restar los coeficientes re-escalados del estudio de evento para el proxy (mostrado en el panel d) de la figura 3).

```
. xteventplot, y ytitle("Coefficient") xtitle("Event time")
 . xteventplot, proxy ytitle("Coefficient") xtitle("Event time")
 . xteventplot, overlay(iv) ytitle("Coefficient") xtitle("Event time")
 . xteventplot, ytitle("Coefficient") xtitle("Event time")
```

Estimación con efectos heterogéneos por cohorte

`xtevent` permite la estimación en entornos con efectos heterogéneos que varían según la cohorte utilizando el estimador de Sun y Abraham (2021) a través de las opciones `cohort` y `control_cohort`.

En la opción `cohort`, el usuario debe especificar la variable que identifica las cohortes, y en la opción `control_cohort` , la variable que identifica la cohorte a utilizar como grupo de control. En el siguiente ejemplo, primero se crea la variable `time_of_treat` para identificar las cohortes y luego se crea la variable `last_treat` para indicar el grupo de control, que en este caso son las últimas unidades tratadas.

```
. gen timet=t if z==1
(1,059 missing values generated)

. by id: egen time_of_treat=min(timet)

. gen last_treat=time_of_treat==39

. replace time_of_treat = . if last_treat
(80 real changes made, 80 to missing)

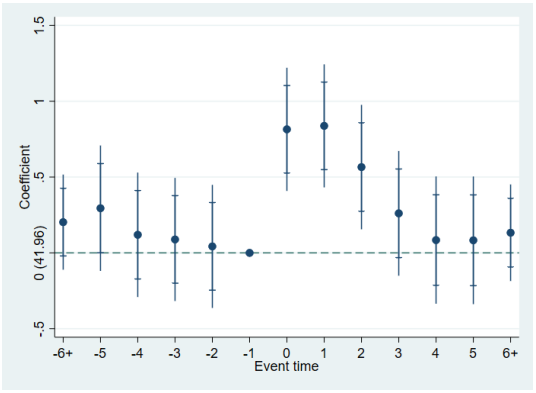
. xtevent y, panelvar(id) timevar(t) policyvar(z) window(5) impute(nuchange) ///
> cohort(variable time_of_treat) control_cohort(variable last_treat)

No proxy or instruments provided. Implementing OLS estimator

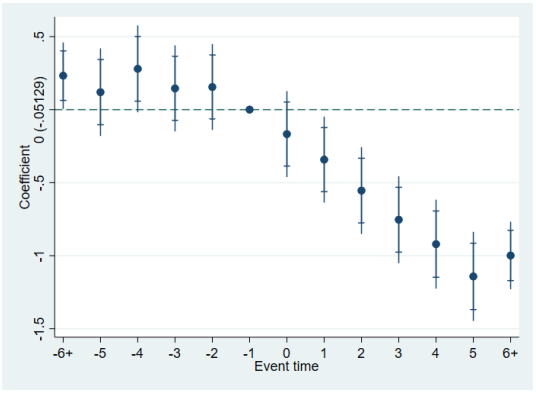
You have specified the cohort or the sunabraham option
Event-time coefficients will be estimated with the Interaction Weighted Estimator
> of Sun and Abraham (2021)

Linear regression, absorbing indicators             Number of obs     =  2,000
Absorbed variable: id                               No. of categories =     50
                                                    F(329, 1621)      =   1.33
                                                    Prob > F          = 0.0003
                                                    R-squared         = 0.2347
```

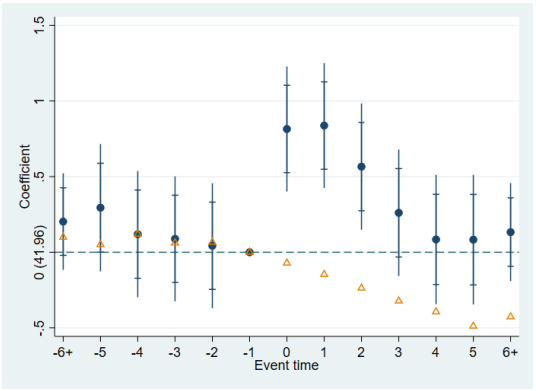
Figura 3: Ajuste de las tendencias previas utilizando variables proxy para la variable de confusión de acuerdo con Freyaldenhoven et al. (2019)



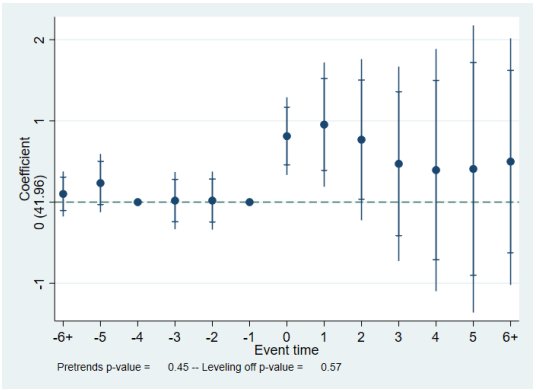
(a) Gráfico de estudio de eventos para el resultado



(b) Gráfico de estudio de eventos para el proxy



(c) Alinear el proxy con el resultado



(d) Restar del resultado la variable de confusión re-escalada

Fuente: Elaborado por los autores con datos de Freyaldenhoven et al. (2022).

```
                                    Adj R-squared     = 0.0562
                                    Root MSE          = 0.7158
```

y	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
_k_eq_m6	.1688391	.1394638	1.21	0.226	-.1047091	.4423872
_k_eq_m5	.3559271	.2014006	1.77	0.077	-.0391058	.7509601
_k_eq_m4	.1412944	.1802456	0.78	0.433	-.2122445	.4948332
_k_eq_m3	.1332897	.1908158	0.70	0.485	-.2409819	.5075613
_k_eq_m2	.1268691	.1961503	0.65	0.518	-.2578657	.5116038
_k_eq_p0	.8289953	.1805352	4.59	0.000	.4748883	1.183102
_k_eq_p1	.8549653	.1870247	4.57	0.000	.4881298	1.221801
_k_eq_p2	.6015642	.1858662	3.24	0.001	.237001	.9661274
_k_eq_p3	.287782	.1883399	1.53	0.127	-.0816333	.6571973
_k_eq_p4	.090265	.185755	0.49	0.627	-.2740801	.45461

```
    _k_eq_p5 |    .090156   .2027774     0.44   0.657    -.3075774    .4878894
    _k_eq_p6 |   .2033672   .1394957     1.46   0.145    -.0702436    .4769781
-------------+----------------------------------------------------------------
```

La salida es análoga a la salida estándar de `xtevent`. La trayectoria de tiempo de evento estimada corresponde a la media ponderada de las estimaciones de estudio de evento comparando cada cohorte de tratamiento con la cohorte de control. `xtevent` guarda las estimaciones por cohorte en las matrices `e(b_interact)` y `e(V_interact)`.

Trayectoria más suave del factor de confusión consistente con las estimaciones

`xteventplot` permite la estimación y visualización de la trayectoria más suave (least wiggly path) del factor de confusión que sea consistente con las estimaciones discutidas en la sección 2, a través de la opción `smpath([type, subopt])`. El tipo de gráfico por defecto es `line`. Con las subopciones `maxorder`, `maxiter`, y `technique`, se puede controlar el orden máximo del polinomio para la trayectoria del factor de confusión y el proceso de optimización.

```
. qui xtevent y x , panelvar(id) timevar(t) policyvar(z) window(5) ///
>   impute(nuchange)

. xteventplot, ytitle("Coefficient") xtitle("Event time") ///
> smpath(line, maxorder(9) maxiter(200) technique(nr 10 dfp 10))

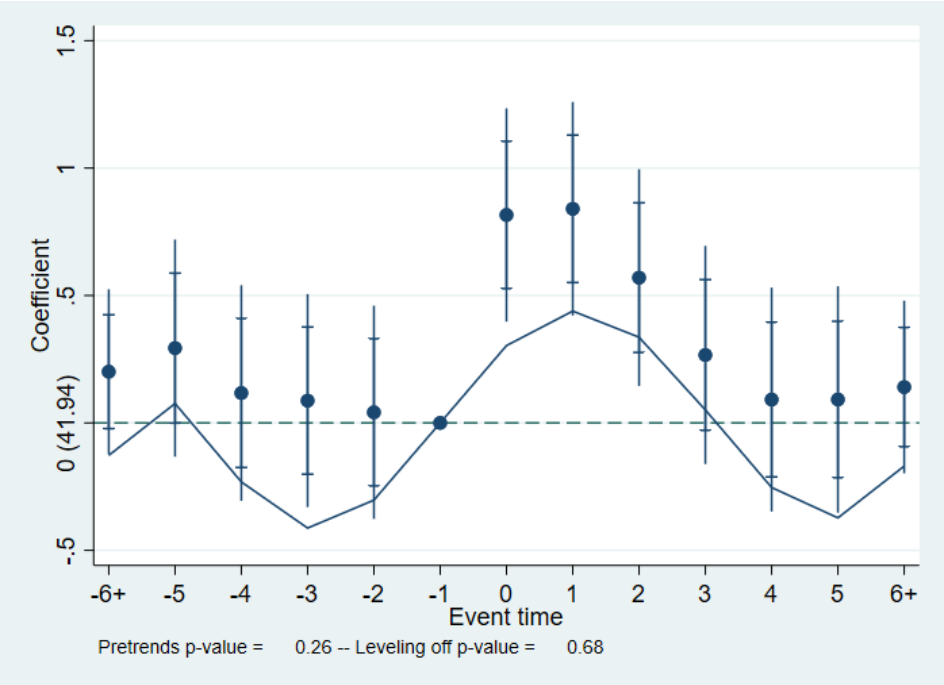
Note: Smoothest line drawn for system confidence level = 95%
Wald Critical Value 21.0260698
Order 0
Wald value 99.1427915
Order 1
Wald value 99.1021008
Order 2
Wald value 77.664952
Order 3
Wald value 64.4768055
Order 4
Wald value 50.446398
Order 5
Wald value 26.7712056
Order 6
Wald value 19.9659905
(setting technique to nr)
Iteration 0:  f(p) =  42803.512  (not concave)
  (output omitted)
```

En este caso, el orden mínimo del polinomio requerido para atravesar la región de confianza es de orden 6. La figura 4 muestra la trayectoria más suave resultante. Si toda la dinámica de la trayectoria estimada en el tiempo del evento de la variable de resultado se debiera a este factor de confusión no observado, el salto en el momento del

evento implicaría que el factor de confusión también tendría que experimentar un salto en ese mismo momento.

Una trayectoria de confusión con estas características sugiere que los efectos estimados pueden deberse realmente a un efecto de la política y no a un factor de confusión.

Figura 4: Trayectoria más suave a través de la región de confianza



Fuente: Elaborado por los autores con datos de Freyaldenhoven et al. (2022).

3.2. Aplicación empírica

Martínez (2022) analiza el efecto de una reforma fiscal en el cantón suizo de Obwalden en 2006. La reforma modificó el esquema del impuesto sobre la renta, reduciendo la tasa impositiva para los contribuyentes con ingresos altos. Hay que enfocarse en el efecto de la reforma sobre los ingresos fiscales de Obwalden procedentes del impuesto sobre el patrimonio, de acuerdo con la figura 9 de Martínez (2022). Se usan los datos de Martínez (2023).

Los datos son un panel equilibrado de 702 observaciones en 26 cantones entre 1990 y

2016. Sólo se trata un cantón, el de Obwalden. Se estima una versión de la ecuación (2):

$$y_{it} = \sum_{k=-5,k\neq-1}^{9} \delta_k \Delta z_{i,t-k} + \delta_{10} z_{i,t-10} + \delta_{-6}(1 - z_{i,t+6}) + \alpha_i + \gamma_t + \varepsilon_{it}. \tag{7}$$

Aquí, el resultado y_{it} son los ingresos per cápita procedentes de los impuestos sobre el patrimonio en el cantón i y el año t, normalizado a 100 en 2005. La variable de política z_{it} es uno para Obwalden en 2006 o después y cero en caso contrario. Los parámetros δ son estimaciones del efecto acumulativo de la reforma en varios horizontes. Se normaliza $\delta_{-1} = 0$. Los parámetros α_i y γ_t representan efectos fijos de cantón y año, respectivamente, y ε_{it} es un término de error.

Se empieza por cargar la base de datos, especificando las variables de unidad y tiempo, y mostrando algunos valores de las variables dependientes y políticas en torno al año de tratamiento:

```
. use martinez.dta, clear

. xtset cant year

Panel variable: cant (strongly balanced)
 Time variable: year, 1990 to 2016
         Delta: 1 unit

. list cant year pcrev_weatax policyvar if cant==6 & inrange(year,2003,2009),
> ab(19) noo sep(7)
```

cant	year	pcrev_weatax	policyvar
OW	2003	98.06313	0
OW	2004	99.06337	0
OW	2005	100	0
OW	2006	92.90456	1
OW	2007	87.39193	1
OW	2008	64.9511	1
OW	2009	73.20993	1

Ahora se estima la ecuación (7) para captar los efectos dinámicos de la reforma sobre los ingresos fiscales. Para ajustar las tendencias previas, se utiliza un ajuste de tendencia lineal basado en los cinco períodos inmediatamente anteriores al evento. No es necesario especificar las variables de unidad y tiempo, ya que los datos fueronse usó `xtset` previamente. La opción `reghdfe`, permite estimar esta ecuación utilizando el comando `reghdfe` de Correia (2016). El uso de `reghdfe` permite errores estándar agrupados bidireccionales a través de la opción `vce`. También se especifica la regla de imputación `stag` y se añaden ponderaciones.

```
. xtevent pcrev_weatax [aweight = weight_pcrev_weatax], panelvar(cant) ///
> timevar(year) policyvar(policyvar) window(-5 9) impute(stag) reghdfe ///
> vce(cluster cant) trend(-5, method(ols))

No proxy or instruments provided. Implementing OLS estimator
(MWFE estimator converged in 3 iterations)

HDFE Linear regression                            Number of obs   =        702
Absorbing 2 HDFE groups                           F(  13,     25) =      98.29
Statistics robust to heteroskedasticity           Prob > F        =     0.0000
                                                  R-squared       =     0.8426
                                                  Adj R-squared   =     0.8265
                                                  Within R-sq.    =     0.0043
Number of clusters (cant)    =         26         Root MSE        =    30.0172

                                (Std. err. adjusted for 26 clusters in cant)
```

pcrev_weatax	Coefficient	Robust std. err.	t	P>|t|	[95% conf.	interval]
_k_eq_m6	19.42696	9.159563	2.12	0.044	.5624837	38.29143
_k_eq_p0	-21.213	5.02208	-4.22	0.000	-31.55617	-10.86983
_k_eq_p1	-40.46391	7.878604	-5.14	0.000	-56.6902	-24.23762
_k_eq_p2	-69.30378	11.09302	-6.25	0.000	-92.15027	-46.45728
_k_eq_p3	-65.53987	15.00997	-4.37	0.000	-96.45348	-34.62627
_k_eq_p4	-55.14003	17.70663	-3.11	0.005	-91.60752	-18.67253
_k_eq_p5	-57.77175	21.48921	-2.69	0.013	-102.0296	-13.51388
_k_eq_p6	-51.15706	21.70713	-2.36	0.027	-95.86374	-6.450375
_k_eq_p7	-43.37057	24.54034	-1.77	0.089	-93.91234	7.17121
_k_eq_p8	-54.18095	29.12688	-1.86	0.075	-114.1689	5.806996
_k_eq_p9	-53.77577	29.91591	-1.80	0.084	-115.3887	7.8372
_k_eq_p10	-44.05725	14.69578	-3.00	0.006	-74.32377	-13.79073
_ttrend	2.528577	2.67721	0.94	0.354	-2.985241	8.042394
_cons	123.7051	9.13919	13.54	0.000	104.8826	142.5276

```
Absorbed degrees of freedom:
```

Absorbed FE	Categories	- Redundant	= Num. Coefs	
cant	26	26	0	*
year	27	0	27	

```
* = FE nested within cluster; treated as redundant for DoF computation
```

Se usó `xteventplot` para mostrar un gráfico de estudio de eventos. `xtevent` permite varias opciones para modificar la apariencia del gráfico. El gráfico se modificó para suprimir las bandas de confianza sup-t y los valores p de las pruebas de sobreidentificación. También se cambiaron los colores y se añadieron títulos a los ejes.

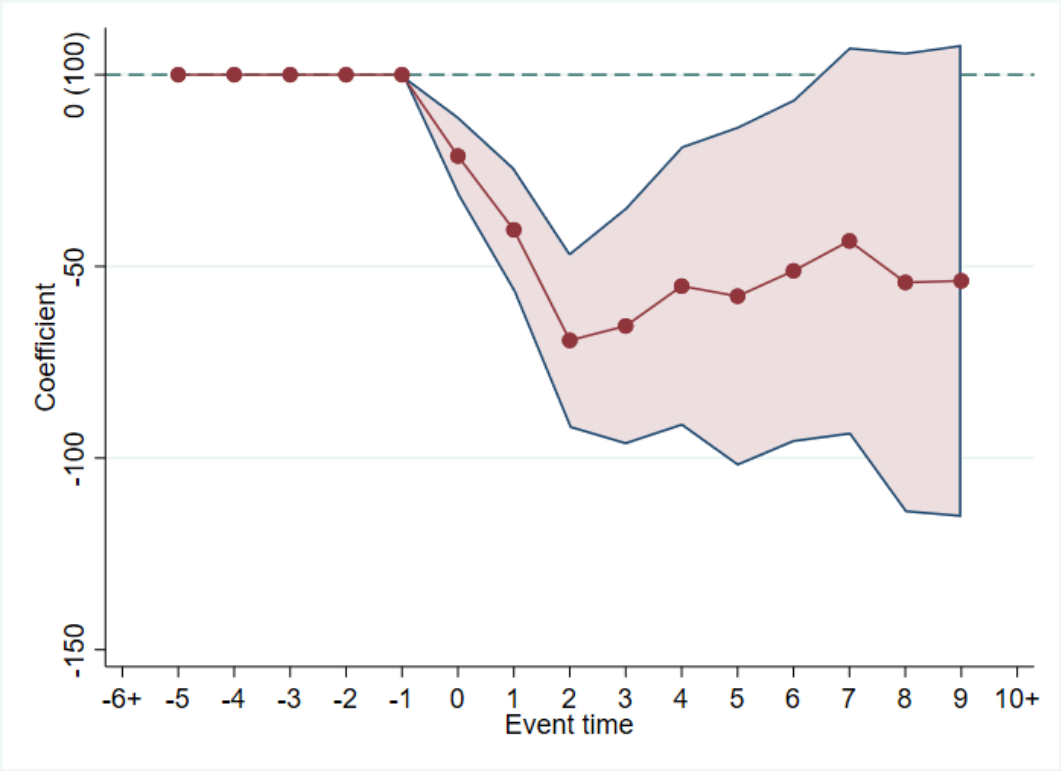
```
. xteventplot, ytitle("Coefficient") xtitle("Event time") ///
> nosupt noprepval nopostpval ///
> scatterplotopts(lcolor(maroon) recast(connected) mcolor(maroon) msymbol(circle))
>  ///
> ciplotopts(recast(rarea) fcolor(maroon*0.2)) ///
> graphregion(fcolor(white))
```

```
option nosupt has been specified. Sup-t confidence bands won´t be displayed or cal
> culated

option noprepval has been specified. The p-value for a pretrends test won´t be dis
> played

option nopostpval has been specified. The p-value for a test of effects leveling-o
> ff won´t be displayed
```

Figura 5: Efecto dinámico de una reforma fiscal suiza según Martínez (2022)



Fuente: Elaboración de los autores con datos de Martínez (2022).

La figura 5 indica que la introducción de la reforma fiscal regresiva en Obwalden redujo los ingresos fiscales de su gobierno (medidos per cápita y en relación con el nivel de 2005). El descenso más importante fue del 69 % y se produjo dos años después de la reforma.

Pruebas de hipótesis

`xteventtest` se puede utilizar para probar diferentes hipótesis sobre los coeficientes de estudio de evento después de `xtevent`. Por ejemplo, la estimación se repite con

errores estándar agrupados por cantón y se utiliza la opción `coefs` para probar si los coeficientes para los tiempos de evento 0, 1, 2 y 3 son iguales a cero conjuntamente.

```
. xteventtest, coefs(0 1 2 3)

 ( 1)  _k_eq_p0 = 0
 ( 2)  _k_eq_p1 = 0
 ( 3)  _k_eq_p2 = 0
 ( 4)  _k_eq_p3 = 0

       F(  4,    25) =   22.82
            Prob > F =    0.0000
```

La prueba indica que los efectos son diferentes de cero. También se puede comprobar si los efectos estimados de las políticas son constantes a lo largo del tiempo:

```
. xteventtest, constanteff

Test for constant post-event coefficients

 ( 1)  _k_eq_p0 - _k_eq_p1 = 0
 ( 2)  _k_eq_p0 - _k_eq_p2 = 0
 ( 3)  _k_eq_p0 - _k_eq_p3 = 0
 ( 4)  _k_eq_p0 - _k_eq_p4 = 0
 ( 5)  _k_eq_p0 - _k_eq_p5 = 0
 ( 6)  _k_eq_p0 - _k_eq_p6 = 0
 ( 7)  _k_eq_p0 - _k_eq_p7 = 0
 ( 8)  _k_eq_p0 - _k_eq_p8 = 0
 ( 9)  _k_eq_p0 - _k_eq_p9 = 0

       F(  9,    25) =   73.46
            Prob > F =    0.0000
```

Esta prueba sugiere que los efectos no son constantes a lo largo del tiempo. Por último, se realizó una prueba de sobreidentificación para ver si los efectos se estabilizan. Para ello, se le pide a `xteventtest` que utilice los dos últimos coeficientes posteriores al evento.

```
. xteventtest, overidpost(2)

Overidentification test for effects leveling off: 2 last post-event coefficients a
> re equal

 ( 1)  - _k_eq_p8 + _k_eq_p9 = 0

       F(  1,    25) =    0.01
            Prob > F =    0.9086
```

Esta prueba sugiere que los efectos se nivelan y que el número de periodos posteriores al evento en el modelo puede ser suficiente para captar los efectos dinámicos de la política.

4. Conclusiones

Este análisis presenta el comando `xtevent` para estimar modelos lineales de panel de datos y visualizar los resultados en gráficos de estudio de eventos. El paquete permite la estimación y el trazado con variables de políticas generales, dando cabida a entornos como serían los variables de políticas continuas y tratamientos reversibles. La versatilidad del paquete y los enfoques de estimación que se sugieren en el análisis se basan en los supuestos que subyacen a la estructura de la ecuación (1), como el hecho de que la variable de política tiene efectos homogéneos y lineales que se pueden separar de los de la variable de confusión. Existen otros enfoques en la literatura, como el de Callaway et al. (2024), para estimar los efectos del tratamiento en entornos con tratamientos continuos y en los que la ecuación (1) puede no cumplirse. El desarrollo de herramientas para aplicar estos enfoques y estimar y visualizar los efectos de las políticas en contextos generales al tiempo que se relajan los supuestos paramétricos como los de la ecuación (1) es un área potencial para futuras investigaciones.

Referencias

Amemiya, Takeshi (1978), "A note on a random coefficients model." *International Economic Review*, 793–796.

Athey, Susan y Guido W Imbens (2022), "Design-based analysis in difference-in-differences settings with staggered adoption." *Journal of Econometrics*, 226, 62–79.

Borusyak, Kirill (2023), "did_imputation and event_plot." `https://github.com/borusyak/did{_}imputation`, consultado en septiembre 2023.

Burkhauser, Richard V, Drew McNichols, y Joseph J Sabia (2023), "Minimum wages and poverty: New evidence from dynamic difference-in-differences estimates." Technical report, National Bureau of Economic Research.

Busch, Alexander y Daniele Girardi (2023), "LPDID: Stata module implementing Local Projections Difference-in-Differences (LP-DiD) estimator." Statistical Software Components, Boston College Department of Economics, URL `https://ideas.repec.org/c/boc/bocode/s459273.html`.

Butts, Kyle (2023), "did2s." `https://github.com/kylebutts/did2s_stata`, consultado en marzo 2024.

Callaway, Brantly, Andrew Goodman-Bacon, y Pedro H. C Sant'Anna (2024), "Difference-in-differences with a continuous treatment." Working Paper 32117, National Bureau of Economic Research.

Callaway, Brantly y Pedro HC Sant'Anna (2021), "Difference-in-differences with multiple time periods." *Journal of Econometrics*, 225, 200–230.

Clarke, Damian y Kathya Tapia-Schythe (2021), "Implementing the panel event study." *The Stata Journal*, 21, 853–884, URL `https://doi.org/10.1177/1536867X211063144`.

Correia, Sergio (2016), "Linear models with high-dimensional fixed effects: An efficient and feasible estimator." Technical report. Working Paper.

Correia, Sergio (2019), "Reghdfe: Stata module to perform linear or instrumental-variable regression absorbing any number of high-dimensional fixed effects."

Cáceres Bravo, Mauricio (2023), "staggered." `https://github.com/mcaceresb/stata-staggered`, consultado en marzo 2024.

de Chaisemartin, Clément, Xavier D'Haultfoeuille, y Yannick Guyonvarch (2019), "DID_MULTIPLEGT: Stata module to estimate sharp Difference-in-Difference designs with multiple groups and periods." Statistical Software Components, Boston College Department of Economics, URL `https://ideas.repec.org/c/boc/bocode/s458643.html`.

Dobkin, Carlos, Amy Finkelstein, Raymond Kluender, y Matthew J Notowidigdo (2018), "The economic consequences of hospital admissions." *American Economic Review*, 108, 308–352.

Fenizia, Alessandra y Raffaele Saggio (2024), "Organized crime and economic growth: evidence from municipalities infiltrated by the mafia." Technical report, National Bureau of Economic Research.

Freyaldenhoven, Simon, Christian Hansen, Jorge Pérez Pérez, y Jesse M Shapiro (2022), "Replication data for "visualization, identification, and estimation in the linear panel event-study design"." `https://data.nber.org/data-appendix/w29170/`. Consultado en septiembre 2022.

Freyaldenhoven, Simon, Christian Hansen, Jorge Pérez Pérez, y Jesse M Shapiro (Forthcoming), "Visualization, identification, and estimation in the linear panel event-study design." In *Advances in Economics and Econometrics: Twelfth World Congress* (Victor Chernozhukov, Eliana Hörner, Johannes La Ferrara, y Ivan Werning, eds.), Cambridge University Press, Cambridge, UK.

Freyaldenhoven, Simon, Christian Hansen, Jorge Pérez Pérez, y Jesse M Shapiro (2021), "Visualization, identification, and estimation in the linear panel event-study design." Working Paper 29170, National Bureau of Economic Research.

Freyaldenhoven, Simon, Christian Hansen, y Jesse M Shapiro (2019), "Pre-event trends in the panel event-study design." *American Economic Review*, 109, 3307–38.

Freyberger, Joachim y Yoshiyasu Rai (2018), "Uniform confidence bands: Characterization and optimality." *Journal of Econometrics*, 204, 119–130.

Goodman-Bacon, Andrew (2021), "Difference-in-differences with variation in treatment timing." *Journal of Econometrics*, 225, 254–277.

Guimarães, Paulo y Pedro Portugal (2010), "A simple feasible procedure to fit models with high-dimensional fixed effects." *The Stata Journal*, 10, 628–649.

Gupta, Sumedha, Dario Salcedo, y Kosali I Simon (2024), "School closures and parental mental health." Working Paper 32516, National Bureau of Economic Research, URL `http://www.nber.org/papers/w32516`.

Hansen, Christian B (2007), "Generalized least squares inference in panel and multilevel models with serial correlation and fixed effects." *Journal of econometrics*, 140, 670–694.

Hegland, Thomas A. (2023), "WOOLDID: Stata module to estimate Difference-in-Differences Treatment Effects with Staggered Treatment Onset Using Heterogeneity-Robust Two-Way Fixed Effects Regressions." Statistical Software Components, Boston College Department of Economics, URL `https://ideas.repec.org/c/boc/bocode/s459238.html`.

Hjort, Jonas y Jonas Poulsen (2019), "The arrival of fast internet and employment in africa." *American Economic Review*, 109, 1032–79, URL `https://www.aeaweb.org/articles?id=10.1257/aer.20161385`.

Korovkin, Vasily y Alexey Makarin (2023), "Conflict and intergroup trade: Evidence from the 2014 russia-ukraine crisis." *American economic review*, 113, 34–70.

Martínez, Isabel Z (2022), "Mobility responses to the establishment of a residential tax haven: Evidence from switzerland." *Journal of Urban Economics*, 129, 103441.

Martínez, Isabel Z (2023), "Replication data for "mobility responses to the establishment of a residential tax haven: Evidence from switzerland"." `https://www.dropbox.com/s/3v0fgzy07jmm3xp/`. Consultado en Abril 2023.

Montiel Olea, José Luis y Mikkel Plagborg-Møller (2019), "Simultaneous confidence bands: Theory, implementation, and an application to svars." *Journal of Applied Econometrics*, 34, 1–17.

Rios-Avila, Fernando (2022), "JWDID: Stata module to estimate Difference-in-Difference models using Mundlak approach." Statistical Software Components, Boston College Department of Economics, URL `https://ideas.repec.org/c/boc/bocode/s459114.html`.

Rios-Avila, Fernando, Pedro H.C. Sant'Anna, y Brantly Callaway (2021), "CSDID: Stata module for the estimation of Difference-in-Difference models with multiple time periods." Statistical Software Components, Boston College Department of Economics, URL `https://ideas.repec.org/c/boc/bocode/s458976.html`.

Roth, Jonathan (2022), "Pretest with caution: Event-study estimates after testing for parallel trends." *American Economic Review: Insights*, 4, 305–322.

Schmidheiny, Kurt y Sebastian Siegloch (2023), "On event studies and distributed-lags in two-way fixed effects models: Identification, equivalence, and generalization." *Journal of Applied Econometrics*, 38, 695–713.

Sun, Liyang (2023), "Eventstudyinteract." `https://github.com/lsun20/EventStudyInteract`, consultado en septiembre 2023.

Sun, Liyang y Sarah Abraham (2021), "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects." *Journal of Econometrics*, 225, 175–199.

A. El paquete xtevent

El paquete `xtevent` incluye comandos `xtevent` para la estimación, `xteventplot` xteventplot para la visualización y `xteventtest` para la comprobación de hipótesis tras la estimación. También incluye `get_unit_time_effects`, un comando auxiliar para utilizar en combinación con xtevent en configuraciones de corte transversales repetidas. Esta sección describe la sintaxis y las opciones de cada uno de estos comandos.

A.1. El comando xtevent

El comando `xtevent` tiene la siguiente sintaxis:

`xtevent` *depvar* [*indepvars*] [*if*] [*in*] [*weight*] , *policyvar(varname)* *panelvar(varname)* *timevar(varname)* [*options*]

A.1.1. Opciones

A.1.2. Principal

`policyvar(`*varname*`)` especifica la variable de política de interés.

`policyvar()` es obligatorio.

`panelvar(`*varname*`)` especifica la variable identificadora transversal que identifica los panels. `panelvar()` es obligatorio si los datos no se han sometido a `xtset` previamente. Véase `xtset`.

`timevar(`*varname*`)` especifica la variable de tiempo. `timevar()` es necesaria si los datos no han sido sometidos previamente a `xtset`. Véase `xtset`.

`window(`*numlist*`)` especifica la ventana alrededor del evento de cambio de política dentro de la cual estimar efectos dinámicos.

`window(k)` con un único número entero positivo $k > 0$ utiliza una ventana simétrica de k periodos alrededor del evento. Por ejemplo, si $k = 2$, habrá cinco coeficientes en la ventana $(-2, -1, 0, 1, 2)$ y dos puntos finales: −3 y +3.

`window(k1 k2)` con dos enteros distintos $k_1 \leq 0$ y $k_2 \geq 0$ utiliza una ventana asimétrica con k_1 periodos antes del evento y $k_2$periodos después del mismo. Por ejemplo, con $k_1 = -1$ y $k_2 = 2$, habrá cuatro coeficientes en la ventana $(-1, 0, 1, 2)$ y dos puntos finales: −2 y +3.

`window(max)` utiliza la ventana más grande posible con los tiempos de evento mínimo y máximo en la muestra de estimación, teniendo en cuenta los puntos finales.

`window(max)` sólo se permite si la política sigue la adopción escalonada y requiere `impute(stag)` o `impute(instag)`(véase abajo).

`window(balanced)` utiliza la ventana más grande posible con los tiempos de evento mínimo y máximo en la muestra de estimación para la que todas las unidades transversales tienen datos.

`window(balanced)` sólo se permite si la política sigue una adopción escalonada y requiere que se especifique `impute(stag)` o `impute(instag)` (véase abajo).

`window()` es necesaria a menos que se especifique `static` o si la ventana de estimación se especifica utilizando las opciones `pre(), post(), overidpre()` y `overidpost()` (Ver abajo).

`pre, post, overidpre` y `overidpost` ofrecen una forma alternativa de especificar la ventana de estimación: `pre` es el número de periodos anteriores al evento en los que se permiten efectos de anticipación. Con `window`, `pre` is 0.

`post` es el número de periodos posteriores al evento en los que se permiten los efectos de la política. Con `window`, `post` es el número de periodos posteriores al evento (sin incluir el periodo para el evento, por ejemplo, momento del evento = 0), excepto los dos últimos periodos (asignados a `overidpost` para la prueba de nivelación).

`overidpre` es el número de periodos anteriores al acontecimiento para una prueba de sobreidentificación de tendencias previas. Con `window`, `overidpre` es el número de periodos anteriores al suceso.

`overidpost` es el número de periodos posteriores al evento para una prueba de sobreidentificación de efectos de nivelacion. Con `window`, `overidpost` es 2.

Sólo se puede declarar una de las siguientes opciones: `window` or `pre, post, overidpre` y `overidpost`.

`static` estima un modelo de datos de panel estático y no genera ni traza variables dummy ficticias de tiempos de evento. `static` no está permitido con `window`, `pre`, `post`, `overidpre`, `overidpost`, o `diffavg`.

`impute(`*type*`, [`*saveimp*`])` imputa los valores adelantados, rezagados y ausentes en *policyvar* policyvar y utiliza esta nueva variable como *policyvar* real. `type` determina la regla de imputación. La subopción `saveimp` añade la nueva variable a la base de datos como *policyvar_imputed*. Los siguientes tipos de imputación están disponibles: `impute(nuchange)` imputa los valores que faltan en *policyvar* según *no unobserved*

change: supone que para cada unidad: i) en los periodos anteriores al primer valor observado, el valor de la política es el mismo que el primer valor observado y; ii) en los periodos posteriores al último valor observado, el valor de la política es el mismo que el último valor observado.

`impute(stag)` aplica *no unobserved change* si *policyvar* cumple los supuestos de adopción escalonada para todas las unidades: i) *policyvar* debe ser binaria, y ii) una vez que *policyvar*alcanza el estado de política adoptada, nunca vuelve al estado de política no adoptada. Véase Freyaldenhoven et al. (Forthcoming) para una explicación detallada del caso de adopción escalonada. `impute(instag)` aplica `impute(stag)`y, además, imputa los valores que faltan dentro del intervalo de datos observados: un valor que falta o un grupo de ellos se imputarán sólo si están precedidos y seguidos por el estado de política no adoptada o por el estado de política adoptada. Véase la sección E del Apéndice E para un ejemplo detallado de la opción `impute`.

`norm(`*integer*`)` especifica el coeficiente de tiempo de evento que debe normalizarse a 0. El valor por defecto es normalizar el coeficiente en -1.

`diffavg` calcula la diferencia en promedios entre los coeficientes estimados después del evento y los coeficientes estimados antes del evento. También calcula su error estándar con `lincom`. `diffavg` no está permitido con `static`.

`savek(`*stub* `[,`*subopt*`])` guarda las variables de tiempo para el evento , tiempo de eventos, tendencia e interacciones. Las variables dummy de tiempo de evento se guardan como *stub_eq_m#* para la variable dummy # # periodos antes del cambio de política, y *stub_eq_p#* fpara la variable dummy # periodos después del cambio de política. La variable dummy para el tiempo de cambio de política es *stub_eq_p0*. El tiempo de evento se guarda como *stub_evtime*. La tendencia se guarda como *stub_trend*. Para la estimación con el método de Sun y Abraham (2021) de forma que `cohort` y `control_cohort` o `sunabraham` estén activas, las variables de interacción se guardan como *stub_m#_c#* or *stub_p#_c#*, donde *c#* indica la cohorte. Se pueden especificar las siguientes subopciones: `noestimate` guarda variables dummy de tiempo de evento, tiempos de evento y tendencias sin estimar el modelo. Esta opción es útil si los usuarios desean personalizar sus regresiones y gráficos.

`saveinteract` guarda las variables de interacción si se especifican `cohort` y `control_cohort`, o `sunabraham`. `noestimate` y `saveinteract` no pueden especificarse simultáneamente.

`replace` sustituye las variables de tiempo de evento, la variable de tiempo del evento, tendencia e interacciones a partir de *stub*.

`kvars(stub)` utiliza las variables dummy de tiempo de evento empleadas previamente y guardadas con el prefijo *stub*. Esto puede utilizarse para acelerar la estimación.

`reghdfe` usa `reghdfe` para la estimación, en lugar de `areg, ivregress`, y `xtivreg`. `reghdfe` es útil para bases de datos grandes. Por defecto, absorbe los efectos fijos de panel y los efectos fijos temporales. Para la estimación MCO, la opción `reghdfe` requiere que `reghdfe` y `ftools` estén instalados. Para la estimación IV, también requiere `ivreghdfe` e `ivreg2` esten instalados:

Nótese que los errores estándar pueden diferir y que los clűsters singletones pueden eliminarse utilizando `reghdfe`. Véase Correia (2016).

`addabsorb(`*varlist*`)` especifica los efectos fijos adicionales que deben absorberse cuando se utiliza `reghdfe`. Por defecto, `xtevent` incluye efectos fijos de tiempo y unidad `addabsorb` requiere `reghdfe`.

`plot` muestra un gráfico de estudio de evento por defecto con intervalos de confianza estándar y bandas de confianza sup-t (Montiel Olea y Plagborg-Møller, 2019). Hay opciones adicionales disponibles con el comando de postestimación `xteventplot`.

`nofe` excluye los efectos fijos de panel.

`note` excluye los efectos fijos temporales.

additional_options: opciones adicionales que se pasarán al comando de estimación. Cuando se especifica `proxy` estas opciones se pasan a `ivregress`. Cuando se especifica `reghdfe`estas opciones se pasan a `reghdfe`. De lo contrario, se pasan a `areg` o a `regress` si `nofe` se específica. Esta opción es útil para calcular errores estándar agrupados o para cambiar los informes de regresión.

A.1.3. Estimación de variables instrumentales con variables proxy(Freyaldenhoven et al., Forthcoming)

`proxyiv(`*proxyiv_spec*`)` especifica los instrumentos para la variable proxy de la política. `proxyiv()` admite tres sintaxis para utilizar adelantos de la variable de política o variables adicionales como instrumentos. `proxy` no está permitido con `cohort`, `control_cohort` o `sunabraham`.

`proxyiv(select)` selecciona el adelanto principal con la primera etapa más fuerte entre todos los posibles adelantos de la variable de política diferenciada que se utilizará como instrumento. `proxyiv(select)` es el valor por defecto para el caso

de un proxy principal, un instrumento, y sólo está disponible en este caso.

`proxyiv(# ...)` especifica una *numlist* con los adelantos de la variable de política diferenciada como instrumentos. Por ejemplo, `proxyiv(1 2)` especifica que los dos primeros adelantos de la diferencia de la variable de política se utilizarán como instrumentos.

`proxyiv(`*varlist*`)` especifica una *varlist* con las variables adicionales que se utilizarán como instrumentos.

A.1.4. Control de las tendencias en tiempo de eventos

`trend(`*#1* `[,`*subopt*`])` extrapola una tendencia lineal utilizando los periodos desde el periodo #1 antes del cambio de política hasta un periodo antes del cambio de política, como en Dobkin et al. (2018). Por ejemplo, `trend(-3)` utiliza los coeficientes de los tiempos de evento -3, -2 y -1 para estimar la tendencia. El efecto estimado de la política corresponde a la desviación de la tendencia lineal extrapolada *#1* debe ser menor que −1. `trend` sólo está disponible cuando el coeficiente normalizado es -1 `pre` = 0. Se pueden pasar las siguientes opciones adicionales como subopciones:

`method(`*string*`)` establece el método para estimar la tendencia lineal. Puede ser por mínimos cuadrados ordinarios (*ols*) o método generalizado de los momentos (*gmm*). (*ols*) omite las variables dummies de tiempo de evento de trend (*#1*) a *-1* y añade una tendencia lineal (_ttrend) a la regresión. (*gmm*)utiliza el GMM para calcular la tendencia de los coeficientes dummy del tiempo de evento. El valor por defecto es `method(`*gmm*`)`.

Nótese que los coeficientes para el tiempo de evento negativos diferirán entre `method(ols)` y `method(gmm)`. `method(ols)` omite los coeficientes de tiempo de evento utilizados para calcular la tendencia, mientras que `method(gmm)` los expresa como diferencias de la tendencia lineal estimada.

`saveoverlay` sguarda las estimaciones para el gráfico de superposición producido por `xteventplot, overlay(`*trend*`)`.

A.1.5. Efectos heterogéneos del tratamiento(Sun y Abraham, 2021)

`cohort(`*cohort_spec*`)` especifica cómo identificar las cohortes de tratamiento utilizadas para la estimación de efectos heterogéneos por cohorte utilizando el estimador de Sun y Abraham (2021). `cohort` requiere el módulo `avar` de Stata.

`cohort(variable` *varname*`,[,force])` especifica que la variable categórica `varname` identifica cada cohorte de tratamiento. Por defecto, `xtevent` comprueba la coherencia de la variable cohorte y la variable de política. `force` fuerza a `xtevent` a saltarse esta comprobación. Esto puede ser útil cuando se estiman efectos de tratamiento heterogéneos entre grupos no definidos por cohortes de tratamiento.

`cohort(create,[,save replace])` pide a `xtevent` que cree la variable categórica de cohorte de tratamiento basándose en los valores de la variable de política. `save` añade la nueva variable de cohorte a la base de datos como *policyvar_cohort*. `replace` reemplaza la variable de cohorte si ya existe. La creación automática de la variable de cohorte sólo está disponible en el caso de adopción escalonada.

`control_cohort(control_cohort_spec)`) especifica cómo identificar la cohorte de control utilizada para la estimación de efectos heterogéneos por cohorte utilizando el estimador de Sun y Abraham (2021). `control_cohort` requiere que se especifique `cohort`. `control_cohort` no se permite con `proxy` o `proxyiv`.

`control_cohort(variable` *varname*`,[,force])` especifica que la variable binaria *varname* iidentifica la cohorte de control. Por defecto, `xtevent` ccomprueba la coherencia de la variable cohorte de control y la variable de política. `force` fuerza a `xtevent` a saltarse esta comprobación. Esto puede ser útil cuando se estiman efectos de tratamiento heterogéneos entre grupos no definidos por cohortes de tratamiento.

`control_cohort(create,[,save replace])` pide a `xtevent` que cree la variable binaria de cohorte de control basándose en los valores faltantes de la variable de cohorte. `save` añade la nueva variable de cohorte de control a la base de datos como `policyvar_control_cohort`. `replace`replace reemplaza la variable de cohorte de control si ya existe. `control_cohort(create)` es el valor por defecto si se especifica `cohort(create)` pero no se especifica `control_cohort`.

`sunabraham` es una abreviatura para especificar la estimación con efectos de tratamiento heterogéneos por cohorte utilizando el estimador de Sun y Abraham (2021). `sunabraham` es equivalente a `cohort(create)` y `control_cohort(create)`.

A.1.6. Estimación con datos de corte transversales repetidos

`repeatedcs` indica que la base de datos en memoria es de corte transversal repetid . En este caso, `panelvar` debe indicar los grupos en los que cambia `policyvar`. Por ejemplo, `panelvar` podría indicar los estados en los que cambia `policyvar`

mientras que las observaciones de la base de datos son individuos en cada estado. Un método alternativo para estimar el estudio de eventos en una base de datos de corte transversales repetidos consiste en utilizar primero `get_unit_time_effects` y a continuación `xtevent`. Véase la descripción del comando`get_unit_time_effects` más adelante. Para la estimación de efectos fijos, `repeatedcs` activa `reghdfe`.

A.1.7. Resultados guardados

`xtevent` guarda los siguientes en `e()`:

Escalares

`e(lwindow)`	dejar punto final para ventana de estimación
`e(rwindow)`	punto final correcto para ventana de estimación

Macros

`e(names)`	nombres de las variables para dummies de tiempo de evento
`e(y1)`	media de variable dependiente en tiempo de evento = -1
`e(x1)`	media de variable proxy en el tiempo de evento = -1, cuando solo se especifica un proxy
`e(trend)`	"tendencia" si la estimación incluye extrapolación de una tendencia lineal
`e(trendmethod)`	método usado para estimar la tendencia lineal: puede ser "ols" o "gmm"
`e(cmd)`	comando de estimación: puede ser "regress", "areg", "ivregress", "xtivreg", o "reghdfe"
`e(df)`	grados de libertad
`e(komit)`	lista de rezagos/adelantos omitida de la regresión
`e(kmiss)`	lista de rezagos/adelantos a omitir en el gráfico
`e(ambiguous)`	lista de unidades de corte transversal omitida por tiempos de evento ambiguos
`e(method)`	"ols" o "iv"
`e(cmd2)`	"xtevent"
`e(depvar)`	variable dependiente
`e(pre)`	número de periodos con efectos de anticipación
`e(post)`	número de periodos con efectos de política
`e(overidpre)`	número de periodos a probar para pretendencias
`e(overidpost)`	número de periodos a probar para nivelación de efectos
`e(stub)`	prefijo para variables dummy de tiempo de evento guardadas

Matrices

`e(b)`	vector de coeficientes
`e(V)`	matriz de varianza-covarianza
`e(delta)`	vector de coeficientes de dummies de tiempo de evento
`e(Vdelta)`	matriz de varianza-covarianza de los coeficientes de dummies de tiempo de evento
`e(deltax)`	coeficientes para proxy de estudio de evento que se usará en gráfico overlay
`e(deltaxsc)`	coeficientes escalados para proxy de estudio de evento que se usará en gráfico overlay
`e(deltaov)`	coeficientes para estudio de evento que se usará en gráfico overlay
`e(Vdeltax)`	matriz de varianza-covarianza de coeficientes para proxy de estudio de evento para gráfico overlay
`e(Vdeltaov)`	matriz de varianza-covarianza de coeficientes de estudio de evento para gráfico overlay
`e(mattrendy)`	matriz con valores de tendencia del eje y para gráfico overlay, solo cuando se especifica la tendencia `trend(#1)`
`e(mattrendx)`	matriz con valores de tendencia del eje x para gráfico overlay solo cuando se especifica la tendencia `trend(#1)`.

`e(b_ir)`	cada vector de columnas contiene estimaciones de cada interacción cohorte-tiempo-relativo y controles incluidos en la regresión de interacción. Las variables de interacción se denominan *_interact_m#_c#* o *_interact_p#_c#*, donde *m#* indica *#* períodos antes del cambio de política, *p#* indica*#* períodos después del cambio de política, y *c#* indica la cohorte. Disponible sólo cuando se especifican `cohort` y `control_cohort`, o `sunabraham`
`e(V_ir)`	matriz de covarianza de las interacciones y controles cohorte-tiempo-relativo-relative-time incluidos en la regresión de interacción. Las variables de interacción se denominan *_interact_m#_c#* o *_interact_p#_c#*, donde *m#* indica*#* períodos antes del cambio de política, *p#* indica *#* períodos después del cambio de política, y *c#* indica ta cohorte. Disponible sólo cuando se especifican `cohort` y `control_cohort`, o `sunabraham`.
`e(b_interact)`	cada vector de columnas contiene estimaciones de efectos específicos de cohorte para el tiempo relativo dado, sólo cuando se especifican `cohort` y `control_cohort`, o `opt sunabraham`
`e(V_interact)`	cada vector columna contiene la estimación de la varianza del estimador efecto específico de la cohorte para el tiempo relativo dado, sólo cuando se especifican cohorte y control `cohort` y `control_cohort`, o `sunabraham`
`e(ff_w)`	cada vector columna contiene estimaciones de las proporciones de cohorte subyacentes al tiempo relativo dado, sólo cuando se especifican `cohort` y `control_cohort`, o `sunabraham`.
`e(Sigma_ff)`	estimación de la varianza de los estimadores de la proporción de cohortes, sólo cuando se especifica `cohort` y `control_cohort`, o `sunabraham`

Functions

`e(sample)`	marca la muestra de estimación

A.2. El comando xteventplot

El comando `xteventplot` produce gráficos de estudio de eventos después de `xtevent`. . La sintaxis es la siguiente:

`xteventplot,` [*options*]

A.2.1. Opciones

`suptreps(`*integer*`)` especifica el número de repeticiones para calcular las bandas de confianza sup-t de Montiel Olea y Plagborg-Møller (2019) para los efectos dinámicos. El valor por defecto es 10000.

`overlay(`*string*`)` crea gráficos superpuestos para la extrapolación de tendencias, la estimación por variables instrumentales en presencia de tendencias previas, y efectos de política constantes a lo largo del tiempo.

`overlay(trend)` superpone los coeficientes evento-tiempode tiempo de evento para la trayectoria de la variable dependiente y la tendencia lineal extrapolada. `overlay(trend)` sólo está disponible después de `xtevent, trend(, saveoverlay)`.

`overlay(iv)` superpone la trayectoria de los coeficientes de evento-tiempotiempo de deevento de la variable dependiente y la variable proxy utilizada para inferir la tendencia del factor de de la variable de confusión. `overlay(iv)` sólo está disponible después de `xtevent, proxy() proxyiv()`.

`overlay(static)` superpone los coeficientes evento-tiempo de tiempo de evento del modelo estimado y los coeficientes implícitos en un efecto constante de la política a lo largo del tiempo. Estos coeficientes se calculan (i) estimando un modelo en el que la política afecta al resultado de forma contemporánea y su efecto es constante, (ii) obteniendo valores predichos de la variable de resultado a partir de este modelo de efectos constantes y (iii) realizando una regresión de los valores predichos sobre las variables ficticias del tiempo de evento.

`y` crea un gráfico de estudio de eventos de la variable dependiente en la estimación de variables instrumentales. `y` y sólo está disponible después de `xtevent, proxy() proxyiv()`.

`proxy` crea un gráfico de estudio de evento de la variable proxy en la estimación de variables instrumentales. `proxy` sólo está disponible después de `xtevent, proxy() proxyiv()`.

`levels(`*numlist*`)` personaliza el nivel de confianza para los intervalos de confianza en el gráfico de estudio de evento. Por defecto, `xteventplot`dibuja un intervalo de confianza estándar y una banda de confianza sup-t. `levels` permite diferentes niveles de confianza para los intervalos de confianza estándar. Por ejemplo, `levels(`*90 95*`)` extrae interválos de confianza de nivel 90 % y 95 % junto con una banda de confianza sup-t para el nivel de confianza por defecto de Stata.

`smpath([`*type, subopt*`])` muestra la trayectoria "más suave" a través de la región de confianza de Wald de los coeficientes de tiempo de evento. `type` determina el tipo de línea, que puede ser `scatter` o `line`. `smpath` no está permitido con `noci`.

Las siguientes subopciones de `smpath` controlan el proceso de optimización. Debido a la naturaleza del problema de optimización, los mensajes de error de optimización 4 y 5 (derivadas faltantes) u 8 (regiones planas) pueden ser frecuentes. Sin embargo, los resultados aproximados de la optimización deben estar cerca de los resultados que se obtendrían con la convergencia del proceso de optimización. Modificación de estas subopciones de optimización puede mejorar el comportamiento de la optimización.

`postwindow(`*scalar > 0*`)` establece el número de estimaciones de coeficientes posteriores al evento que se utilizarán para calcular la línea más uniforme. El valor por defecto es utilizar todas las estimaciones de la ventana posterior al evento.

`maxiter(`*integer*`)` sestablece el número máximo de iteraciones internas para la optimización. El valor por defecto es 100.

`maxorder(`*integer*`)` establece el orden máximo para la línea polinómica más uniforme.`maxorder` debe estar entre 1 y 10. El valor por defecto es 10.

`technique(`*string*`)`) establece la técnica de optimización para las iteraciones internas del programa cuadrático. Se permiten "nr," "bfgs," "dfp," y combinaciones de estas. Véase `maximize`. Por defecto es"nr 5 bfgs."

`overidpre` cambia los coeficientes probados en la prueba de sobreidentificación de tendencias previas. Por defecto se prueban todos los coeficientes pre-evento `overidpre(`*#1*`)` prueba si los coeficientes para los periodos *#1* más tempranos antes del evento son iguales a *0*, incluyendo los puntos finales. Por ejemplo, con una ventana de 3, `overidpre(2)` comprueba que los coeficientes para los periodos anteriores al suceso -4+ (el punto final) y -3 son conjuntamente iguales *0*. *#1* debe ser mayor que *0*. Véase `xteventtest` a continuación.

`overidpost` cambia los coeficientes a probar para la prueba de sobreidentificación de nivelación. Por defecto, se comprueba que el coeficiente situado más a la derecha y el anterior sean iguales. `overidpost(`*#1*`)` comprueba si los coeficientes de los últimos *#1* periodos posteriores al suceso son iguales entre sí, incluidos los puntos finales. Por ejemplo, con una ventana de 3, `overidpost(3)` comprueba que los coeficientes para de tiempo de evento s 4+ (el punto final), 3 y 2 son iguales entre sí. *#1* debe ser mayor a *1*. Véase el comando `xteventtest` más abajo.

Las siguientes opciones controlan el aspecto del gráfico:

`noci` omite la visualización y el cálculo de la banda de confianza de Wald y sup-t. `noci` anula `suptreps` si se especifica `noci` no está permitido con `smpath`.

`nosupt` omite la visualización y el cálculo de las bandas de confianza sup-t `nosupt` anula `suptreps` si se especifíca.

`nozeroline` omite la visualización de la línea de referencia en *0*. Nótese que se pueden obtener líneas de referencia con diferentes estilos eliminando la línea por defecto con `nozeroline` y añadiendo otras lineas con `yline`. Véase `added_line_options`.

`nonormlabel` suprime la etiqueta del eje vertical para la media de la variable dependiente en el tiempo de evento correspondiente al coeficiente normalizado.

`noprepval` omite la visualización del valor p para una prueba de tendencias previas. Por defecto, se trata de una prueba de Wald para todos los coeficientes pre-evento iguales a *0*, a menos que se especifique `overidpre`.

`nopostpval` omite la visualización del valor p para una prueba de nivelación de efectos.

Por defecto, se trata de una prueba de Wald de que los últimos coeficientes posteriores al evento sean iguales, a menos que se especifique `overidpost`.

`scatterplotopts` especifica las opciones que se pasarán a `scatter` para el gráfico de coeficientes.

`ciplotopts` especifica las opciones que se pasarán a `rcap` fopara el gráfico del intervalo de confianza. Estas opciones se desactivan si se especifica `noci`.

`suptciplotopts` especifica las opciones que se pasarán a `rcap` para el gráfico de banda de confianza sup-t. Estas opciones se desactivan si se especifica `nosupt`.

`smplotopts`especifica las opciones que se pasarán a `line` fpara la trayectoria más uniforme a través del gráfico de la región de confianza. Estas opciones sólo están activas si se especifica `smpath`.

`trendplotopts` especifica las opciones que se pasarán a `line` para el gráfico de superposición de tendencia extrapolada. Estas opciones sólo están activas si se especifica `overlay(`*trend*`)`.

`staticovplotopts` especifica las opciones que se pasarán a `line` para el gráfico de superposición de efectos estáticos. Estas opciones sólo están activas si se especifica `overlay(`*static*`)`.

`addplots` especifica los gráficos adicionales que se superpondrán al gráfico del estudio de eventos.

`textboxoption` especifica las opciones que se deben pasar al cuadro de texto de las pruebas de pretendencia y de nivelación. Estas opciones se desactivan si se especifica `noprepval` y `nopostval`. Véase `textbox_options`.

additional_options: opciones adicionales que se pasarán a `twoway`. Véase `twoway`.

A.3. El comando xteventtest

El comando `xteventtest` realiza pruebas de hipótesis después del comando `xtevent` La sintaxis es la siguiente:

`xteventtest,` [*options*]

A.3.1. Opciones

`coefs(`*numlist*`)` especifica una lista numérica de tiempos de evento que se van a comprobar. Se comprueba que sean iguales a *0* conjuntamente, a menos que se solicite lo

contrario en `testopts()`.

`cumul` solicita una prueba de igualdad a *0* para la suma de cada coeficiente para cada tiempo de evento en `coefs()`.

`allpre` comprueba que todos los coeficientes pre-evento son iguales a *0*. Con `cumul`, comprueba que la suma de todos los coeficientes pre-evento es igual a *0*.

`allpost` comprueba que todos los coeficientes pre-evento son iguales a *0*. Con `cumul`, comprueba que la suma de todos los coeficientes post-eventos es igual a *0*.

`linpretrend` solicita una prueba de especificación para ver si los coeficientes siguen una tendencia lineal antes del evento.

`trend(`*#1*`)` comprueba si existe una tendencia lineal desde el periodo de tiempo *#1* antes del cambio de política. Utiliza `xtevent, trend(`*#1, method(ols)*`)` para estimar la tendencia. *#1* debe ser menor que *-1*.

`constanteff` comprueba que todos los coeficientes posteriores al suceso son iguales.

`overid` comprueba las restricciones de sobreidentificación: una prueba de tendencias previas y una prueba de nivelación de efectos. Los periodos que se comprueban son los utilizados en la llamada `xtevent`.

`overidpre(`*#1*`)` comprueba la restricción de sobreidentificación de tendencias previas. Comprueba que los coeficientes para los periodos *#1* más tempranos antes del evento son iguales a *0*, incluyendo los puntos finales. Por ejemplo, con una ventana de 3, `overidpre(2)` comprueba que los coeficientes para los tiempos de eventos -4+ (el punto final) y -3 son conjuntamente iguales a *0*. El coeficiente *#1* debe ser mayor a *0*.

`overidpost(`*#1*`)` comprueba la nivelación de los efectos de la restricción de sobreidentificación. Comprueba que los coeficientes para los últimos periodos *#1* después del evento son iguales, incluidos los puntos finales. Por ejemplo, con una ventana de 3, `overidpost(3)` comprueba que los coeficientes para los tiempo de evento 4+ (el punto final), 3 y 2 son iguales entre sí. *#1* debe ser mayor que *1*.

`testopts(`*string*`)` especifica las opciones que se pasarán a `test`. Véase `test`.

A.3.2. Resultados guardados

`xteventtest` guarda lo siguiente en `r()`:

Scalars

`r(p)`	valor p de "dos colas"
`r(F)`	estadístico F
`r(df)`	Grados de libertad de las restricciones de la prueba
`r(df_r)`	Grados de libertad residuales
`r(dropped_i)`	Índice de la i-ésima restricción eliminada
`r(chi2)`	Ji-cuadrado
`r(drop)`	*1* 1 si se eliminaron restricciones, *0* en caso contrario

Macros

`r(mtmethod)`	Método de ajuste para pruebas múltiples. Este macro se hereda de la prueba `test`.

Matrices

`r(mtest)`	Resultados de la prueba múltiple. Esta matriz se hereda de la prueba `test`.

A.4. El comando get_unit_time_effects

El comando `get_unit_time_effects` genera efectos de grupo y de tiempo en una base de datos de corte transversal repetido. Produce un archivo de datos de Stata con las variables *panelvar*, *timevar*, y *_unittimeeffects*. La variable *_unittimeeffects* contiene los efectos grupo-tiempo. Hansen (2007) describe un procedimiento de dos pasos para obtener las estimaciones de los coeficientes de las covariables que varian a nivel de grupo en un marco transversal repetido. El procedimiento de dos pasos puede utilizarse para obtener las estimaciones de los coeficientes de un estudio de evento cuando los datos son de corte transversal repetido. `get_unit_time_effects` implementa la primera parte del procedimiento en dos pasos. A continuación, `xtevent` se puede utilizar para la segunda parte del procedimiento para obtener las estimaciones del coeficiente del estudio de eventos. Véase el Apéndice F. La sintaxis del comando `get_unit_time_effects` es la siguiente:

`get_unit_time_effects` *depvar* [*indepvars*] [*if*] [*in*] [*weight*] , ***panelvar(varname)***
timevar(varname) [*options*]

A.4.1. Opciones

`panelvar(`*varname*`)` especifica la variable de grupo. La variable de política debe variar en este nivel de grupo.

`timevar(`*varname*`)` especifica la variable de tiempo.

`saving(`*filename* `[, replace])` especifica el nombre del archivo de datos de Stata para guardar las estimaciones de los efectos de tiempo unitario. Si no se especifica `saving`

el archivo se guarda en el directorio actual con el nombre *unit_time_effects.dta*. La subopción `replace` sobrescribe el archivo de efectos unidad-tiempo.

`nooutput` omite la tabla de regresión.

`clear` sustituye a la base de datos en memoria por el archivo de efectos unidad-tiempo.

B. Detalles sobre la extrapolación de tendencias

El método por defecto para extrapolar una tendencia lineal utiliza GMM y se implementa de la siguiente manera. Sea $T_G \leq L_G$ el número de periodos anteriores a G G utilizados para estimar los parámetros de la tendencia y $T_M \leq M$ el número de periodos "posteriores al evento" en los que la tendencia está activa. Se presume que $f_k \neq 0$ para $k \in [-G - T_G, T_M]$ y 0 en caso contrario. Entonces hay momentos dados por

$$\widehat{\delta}_k - \phi' f_k = 0$$

para $k = -G - T_G, ..., -G - 1$. Sea $\widehat{\delta}_{T_G}$ el T_G-vector que recoge $\widehat{\delta}_k$. Sea H_{T_G} la $T_G \times \dim(\phi)$ matríz cuya j-ésima fila es f_k' para $k = j - 1 - G - T_G$.

Un estimador de distancia mínima $\widehat{\phi}$ de ϕ resuelve

$$\begin{aligned} \widehat{\phi} &= \arg\min_{\phi} \widehat{h}(\phi)' \widehat{W} \widehat{h}(\phi) \\ \widehat{h}(\phi) &= \widehat{\delta}_{T_G} - H_{T_G}\phi. \end{aligned}$$

Resolviendo condiciones de primer orden, se obtiene

$$\begin{aligned} 0 &= -H_L' \widehat{W}(\widehat{\delta}_{T_G} - H_L \widehat{\phi}) \\ \widehat{\phi} &= (H_{T_G}' \widehat{W} H_{T_G})^{-1} H_{T_G}' \widehat{W} \widehat{\delta} \end{aligned}$$

En condiciones de regularidad adecuadas, se tiene que

$$\sqrt{n} \begin{pmatrix} \widehat{\phi} - \phi_0 \\ \widehat{\delta}_{T_G} - \delta_{L0} \end{pmatrix} \to N\left(0, \begin{bmatrix} \Lambda_{T_G} \Omega_{T_G} \Lambda_{T_G}' & \Lambda_{T_G} \Omega_{T_G} \\ \Omega_{T_G} \Lambda_{T_G}' & \Omega_{T_G} \end{bmatrix}\right)$$

donde

$$\Lambda_{T_G} = (H_{T_G}' W H_{T_G})^{-1} H_{T_G}' W$$

y Ω_{T_G} es la varianza asintótica de $\widehat{\delta}_{T_G}$. La matriz de ponderación eficiente factible es $\widehat{W} = \widehat{\Omega}_{T_G}^{-1} \to \Omega_{T_G}^{-1}$, y con $W = \Omega_{T_G}^{-1}$ se tiene que $\Lambda_{T_G} \Omega_{T_G} \Lambda_{T_G}' = (H_{T_G}' \Omega_{T_G}^{-1} H_{T_G})^{-1}$.

B.1. Estimación e inferencia sobre la trayectoria de tiempo de evento ajustada

Ahora sea $\widehat{\delta}$ el vector que contiene toda la trayectoria estimada en el tiempo del evento, de manera que: $\dim\left(\widehat{\delta}\right) = \dim\left(\delta\right) = M + L_M + G + L_G + 2$. Sea H la matriz $\dim\left(\delta\right) \times \dim\left(\phi\right)$ cuya j-ésima fila es f_k', donde $k = j - 2 - G - L_G$. Dada la estimación $\widehat{\phi}$ obtenemos la estimación ajustada $\widehat{\delta}^*$ de la trayectoria en el tiempo del evento mediante:

$$\widehat{\delta}^* = \widehat{\delta} - H\widehat{\phi}.$$

Sea $\Lambda = \begin{bmatrix} \mathbf{0} & \Lambda_{T_G} & \mathbf{0} \end{bmatrix}$, donde $\mathbf{0}$ son matrices conformables de ceros ($\dim\left(\phi\right) \times 1$ y $\dim\left(\phi\right) \times \left(\dim\left(\delta\right) - L_G\right)$, respectivamente), $\widehat{\Lambda}$ sea su análogo muestral, e I sea una $\dim\left(\delta\right) \times \dim\left(\delta\right)$ matriz de identidad. Por lo tanto,

$$\widehat{\delta}^* = \widehat{\delta} - H\widehat{\phi} = \left(I - H\widehat{\Lambda}\right)\widehat{\delta}$$

y se sigue que (bajo condiciones adecuadas):

$$\sqrt{n}\left(\widehat{\delta}^* - \delta_0^*\right) \to N\left(0, \Omega - H\Lambda\Omega - \Omega\Lambda'H' + H\Lambda\Omega\Lambda'H'\right)$$

donde Ω es la matriz de varianza asintótica de $\hat{\delta}$ y

$$\delta_{0,k}^* = \begin{cases} 0, & k < -G \\ \sum_{m=-G}^{k} \beta_m, & -G \leq k \leq M \\ \sum_{m=-G}^{M} \beta_m, & k > M. \end{cases}$$

Las pruebas de hipótesis para las tendencias previas (pre-trends) y la estabilización de los efectos dinámicos (leveling-off) ahora pueden proceder como en el caso de efectos fijos bidireccionales (TWFE), reemplazando $\widehat{\delta}$ con $\widehat{\delta}^*$.

B.2. Covarianza de la trayectoria de tiempo de evento ajustada y coeficientes sobre controles

Para ciertos propósitos, puede ser interesante probar hipótesis conjuntas sobre (δ_0^*, ψ_0). Dado que $\widehat{\delta}^* = (I - H\widehat{\Lambda})\widehat{\delta}$, se tiene que,

$$\sqrt{n}\begin{pmatrix}\widehat{\delta}^* - \delta_0^* \\ \widehat{\psi} - \psi_0\end{pmatrix} = \begin{pmatrix} I - H\widehat{\Lambda} & 0 \\ 0 & I\end{pmatrix}\sqrt{n}\begin{pmatrix}\widehat{\delta} - \delta_0 \\ \widehat{\psi} - \psi_0\end{pmatrix} \to \begin{pmatrix} I - H\Lambda & 0 \\ 0 & I\end{pmatrix} N(0, V),$$

con 0 matrices conformables de ceros ($\dim(\delta) \times \dim(\psi)$ para la superior derecha y $\dim(\psi) \times \dim(\delta)$ para la parte inferior izquierda) I matrices conformables de identidad ($\dim(\delta)$ para la parte superior izquierda $\dim(\psi)$ para la parte inferior derecha). Por último, pérmitase que Ω_ψ denote la matriz de varianza asintótica de $\widehat{\psi}$ de dimensión $\dim(\psi) \times \dim(\psi)$; y que $\Omega_{\delta,\psi}$ denote la matriz de covarianza asintótica entre $\widehat{\delta}, \widehat{\psi}$ de dimensión $\dim(\psi) \times \dim(\delta)$. Se puede expresar V como

$$V = \begin{pmatrix} \Omega & \Omega_{\delta,\psi}' \\ \Omega_{\delta,\psi} & \Omega_\psi \end{pmatrix}$$

y la varianza asintótica de $\left(\widehat{\delta}^*, \widehat{\psi}\right)$ es

$$\begin{pmatrix} I - H\Lambda & 0 \\ 0 & I\end{pmatrix}\begin{pmatrix} \Omega & \Omega_{\delta,\psi}' \\ \Omega_{\delta,\psi} & \Omega_\psi \end{pmatrix}\begin{pmatrix} I - \Lambda' H' & 0 \\ 0 & I\end{pmatrix} = \begin{pmatrix} (I - H\Lambda)\Omega(I - \Lambda' H') & (I - H\Lambda)\Omega_{\delta,\psi}' \\ \Omega_{\delta,\psi}(I - \Lambda' H') & \Omega_\psi \end{pmatrix}.$$

C. Bandas de confianza Sup-t

Se usan bandas sup-t para la inferencia uniforme (véase Freyberger y Rai (2018), Montiel Olea y Plagborg-Møller (2019), y las referencias de las mismas para antecedentes adicionales). Estas bandas se construyen sumando (o restando) una constante por el vector de errores estándar de $\hat{\delta}$ de $\hat{\delta}$, tal que la banda de confianza simultánea en cada coeficiente δ_k es igual a:

$$\hat{B}_k(\alpha) \equiv \left[\hat{\delta}_k - c_\alpha \hat{\sigma}_k, \hat{\delta}_k + c_\alpha \hat{\sigma}_k\right]$$

para un nivel de significación α elegido, donde c_α denota el valor crítico sup-t correspondiente.

Para calcular c_α, se utiliza un sencillo estimador complementario (Montiel Olea y Plagborg-Møller (2019)).

1. Dibújense N i.i.d vectores $\hat{V}^{(\ell)},\ \ell = 1, \ldots, N$ de dimensión $K = \dim(\hat{\delta})$ a partir de un modelo multivariante normal con media $\mathbf{0}_K$ y varianca $\hat{\Omega}$ dada por la varianza estimada de $\hat{\delta}$.

2. Por cada repetición $\ell = 1, \ldots, N$, let $t_\ell = \underset{k=1,\ldots,K}{\text{máx}} \left| \hat{\Omega}_{kk}^{-1/2} \hat{V}_k^{(\ell)} \right|$.

3. Establézcase $c_\alpha = Q_{1-\alpha}(t_\ell)$, $\ell = 1, \ldots, N$, donde Q Q es la función cuantil.

D. Detalles para la trayectoria más suave (least wiggly path)

D.1. La propuesta de la trayectoria más suave

Denotamos la dimensión de δ como $K \equiv G + M + L_G + L_M + 2$. Para v, un vector de coeficientes de dimensión finita, y k, un entero, definimos el término polinomial:

$$\delta_k^*(v) = \sum_{j=1}^{\dim(v)} v_j \left(\frac{k - s_1}{s_2}\right)^{j-1},$$

donde v_j denota el j -ésimo elemento del vector de coeficientes v y $\dim(v)$ denota la dimensión de este vector. Los valores s_1 y s_2 son constantes que desplazan y escalan el tiempo del evento (el rango del polinomio). Fijamos $s_1 = -G - L_G - 1$ y $s_2 = M + L_M + G + L_G + 2$. Sea $\delta^*(v)$ la colección de elementos $\delta_k^*(v)$ for $-G - L_G - 1 \leq k \leq M + L_M$, de manera que $\delta^*(v)$ refleja una trayectoria polinomial en el tiempo del evento con coeficientes v.

`xtevent` genera gráficos de la trayectoria más "suave" (least wiggly) del factor de confusión cuya trayectoria está contenida en la región de Wald $CR(\delta)$ para la trayectoria en el tiempo del evento del resultado. Específicamente, grafica $\delta^*(v^*)$, donde

$$p^* = \min\{\dim(v) : \delta^*(v) \in CR(\delta)\} \text{ and} \tag{8}$$

$$v^* = \arg\min_v\{v_{p^*}^2 : \dim(v) = p^*, \delta^*(v) \in CR(\delta)\}. \tag{9}$$

Intuitivamente, la región de confianza de Wald representa el conjunto de trayectorias en el tiempo del evento para las cuales una prueba conjunta F de los estimados puntuales observados no es rechazada.Dado que esta región es una elipse, no existe una ilustración gráfica directa de esta región en un gráfico de estudio de evento.

Para graficar la trayectoria más suave (least wiggly path), resolvemos un problema en dos partes: En la ecuación (8), determinamos el menor orden p^* tal que un polinomio de orden p^* esté completamente contenido en la región de Wald $CR(\delta)$. En la ecuación (9), seleccionamos el polinomio con el coeficiente más bajo en el término de mayor orden de dicho polinomio.

En la práctica, normalizamos la trayectoria del evento de manera que $\delta_k = 0$ para al menos un valor de k (por ejemplo, usualmente en $k = -1$).

Usaremos $\mathcal{N}$ para denotar el conjunto de tamaño $|\mathcal{N}|$ que toma todos los coeficientes normalizados, de modo que $\delta_k^*(v) = 0$ para $k \in \mathcal{N}$. A lo largo del análisis, consideramos únicamente el caso $|\mathcal{N}| \in \{1, 2\}$, i.e., es decir, permitimos un máximo de dos normalizaciones

D.2. Implementación

D.2.1. Encontrar p^*

Primero hay que abordar el problema de encontrar p^* en (8). Se define Σ como la matriz de covarianza de $\widehat{\delta}$ con ceros añadidos en las filas y columnas correspondientes a los coeficientes normalizados.

Dado que p^* es generalmente pequeño, es factible resolver (8) iterativamente de la siguiente manera:

Nótese que p^* es menor que $K = \dim(\delta)$ por construcción, y, por lo tanto, está garantizado (al menos teóricamente) que el bucle en el Algoritmo 1 converja después de, como máximo, K iteraciones. Para garantizar la estabilidad numérica, restringimos p^* a que sea menor o igual a diez en nuestra implementación (con una opción para reducir

Algoritmo 1 Encontrar p^*

$p^* \leftarrow 0$
factible $\leftarrow 0$
mientras factible $= 0$ **hacer**
  $p^* \leftarrow p^* + 1$
  factible $\leftarrow$ SolucionEnRegionDeWald$(\widehat{\delta}, p^*, \alpha)$
fin mientras

función SolucionEnRegionDeWald$(\widehat{\delta}, p^*, \alpha)$
  $W^* = \min_{v:dim(v)=p^*}[\delta^*(v) - \widehat{\delta}]'\Sigma^{-1}[\delta^*(v) - \widehat{\delta}]$ s.t. $\delta^*_{k_n}(v) = 0$ para $n \in \mathcal{N}$
  ▷ Σ^{-1} denota la inversa generalizada.
  devolver 1$(W^* \leq c^{1-\alpha})$
  ▷ $c^{1-\alpha}$ es el cuantil $1-\alpha$ de una variable aleatoria $\tau \sim \chi^2(K - |\mathcal{N}|)$.
fin función

aún más este límite superior). Si $p^* > 10$, concluimos que "no existe una trayectoria suave."

Para encontrar W^* en la práctica, utilizamos las condiciones de primer orden de la minimización del estadístico de Wald, sujetas a las restricciones sobre los coeficientes normalizados.

Para lograrlo, escribimos el polinomio de la trayectoria más suave en notación matricial como $\delta^*(v) = \underset{K\times p^* p^*\times 1}{F} v$, donde $F_{kj} = \left(\frac{k-s_1}{s_2}\right)^{j-1}$ para $k = 1, \ldots, K$. Las filas de F recogen los términos del polinomio para un tiempo de evento (desplazado) k, y el vector v recoge los coeficientes del polinomio.

El problema para encontrar W^* se puede reescribir como:

$$\min_v \left[Fv - \widehat{\delta}\right]' \Sigma^{-1} \left[Fv - \widehat{\delta}\right] \text{s.t. } \delta^*_k(v) = 0 \text{ para } k \in \mathcal{N}. \tag{10}$$

A partir del Lagrangiano, las condiciones de primer orden son las siguientes:

$$F'\Sigma^{-1}Fv = F'\Sigma^{-1}\widehat{\delta} + \frac{1}{2}\lambda A'_{norm}$$
$$A_{norm}v = 0$$

Aquí, A_{norm} es la matriz cuyas filas de F corresponden a los coeficientes normalizados. El álgebra muestra entonces que: $v(\lambda) = (F'\Sigma^{-1}F)^{-1}[F'\Sigma^{-1}\widehat{\delta} + \frac{1}{2}\lambda A'_{norm}]$, y al sustituir esta expresión en la segunda condición de primer orden mencionada anteriormente, se obtiene:

$$\lambda = -2[A_{norm}(F'\Sigma^{-1}F)^{-1}A'_{norm}]^{-1}A_{norm}(F'\Sigma^{-1}F)^{-1}F'\Sigma^{-1}\widehat{\delta}.$$

Entonces, la solución para v está dada por

$$\tilde{v} = (F'\Sigma^{-1}F)^{-1} \left[F'\Sigma^{-1}\widehat{\delta} - A'_{norm}[A_{norm}(F'\Sigma^{-1}F)^{-1}A'_{norm}]^{-1}A_{norm}(F'\Sigma^{-1}F)^{-1}F'\Sigma^{-1}\widehat{\delta}\right].$$

Podemos expresar la solución para v como un producto matricial. Sea $XX \equiv \begin{bmatrix} 2F'\Sigma^{-1}F & A'_{norm} \\ A_{norm} & \underset{|\mathcal{N}|\times|\mathcal{N}|}{\mathbf{0}} \end{bmatrix}$ y $Xy \equiv \begin{bmatrix} 2F'\Sigma^{-1}\widehat{\delta} \\ \underset{|\mathcal{N}|\times|\mathcal{N}|}{\mathbf{0}} \end{bmatrix}$. Entonces, la solución para v se obtiene tomando las primeras K filas del vector $\tilde{v} = (XX)^{-1}Xy$.

D.2.2. Encontrar la trayectoria óptima dada p^*

Una vez que se ha encontrado una solución a (8) utilizando el Algoritmo 1, el siguiente paso es encontrar el polinomio con el coeficiente más bajo en el término p^* que todavía está contenido en la region Wald (véase la ecuación 9). Nótese en primer lugar que, por construcción, $v^2_{p^*} \neq 0$ (Si no, Algoritmo 1 habría encontrado una solución en $p^* - 1$). v^* puede hallarse entonces mediante una simple minimización restringida en el vector v (de dimensión p^*):

$$v^* = \arg\min_v v^2_{p^*} \tag{11}$$

$$\text{tal que}[\delta^*(v) - \widehat{\delta}]\Sigma^{-1}[\delta^*(v) - \widehat{\delta}] \leq c^{1-\alpha}\} \tag{12}$$

$$\text{and } \delta^*_k(v) = 0 \text{ para } k \in \mathcal{N}, \tag{13}$$

con Σ y $c^{1-\alpha}$ definidos como arriba.

En primer lugar, si $p^* \leq |\mathcal{N}|$, a restricción en (13) implica que $v^* = 0$ y se ha concluido. A continuación, si $p^* > |\mathcal{N}|$, se observa que v^* siempre estará en el límite de la región de Wald. Así, la restricción en (12) siempre será vinculante, y ambas restricciones se pueden sustituir directamente para resolver para v^*. En particular, dado un conjunto $\mathcal{N}$ de coeficientes normalizados y la restricción en (12), es posible resolver para algunos de los otros coeficientes. Si $p^* > |\mathcal{N}| + 1$, se utiliza una optimización sin restricciones para resolver para los restantes posteriormente.

Específicamente, particionar las matrices A_{norm} y F en tres:

$$A_{norm} = \underset{|\mathcal{N}|\times(p^*-|\mathcal{N}|-1)}{[\ A_b\ }, \underset{|\mathcal{N}|\times|\mathcal{N}|}{A_1\ }, \underset{|\mathcal{N}|\times 1}{A_2\ ]}$$
$$F = \underset{K\times(p^*-|\mathcal{N}|-1)}{[\ F_b\ }, \underset{K\times|\mathcal{N}|}{F_1\ }, \underset{K\times 1}{F_2\ }],$$

con el vector v dividido en $v = [v_b; v_1; v_2]$. Se resuelve para los coeficientes v_1 y v_2 utilizando las restricciones en (13) y (12) respectivamente, y luego se hace lo mismo para los coeficientes v_b por minimización sin restricciones. Primero, notemos que, dado que A_{norm} contiene las filas de F asociados con los coeficientes normalizados, tenemos:

$$A_{norm}v = A_b v_b + A_1 v_1 + A_2 v_2 = 0 \text{ por lo tanto, } v_1 = {A_1}^{-1}(-A_b v_b - A_2 v_2). \tag{14}$$

De ello,

$$\delta^*(v) = Fv = F_b v_b + F_1 v_1 + F_2 v_2 = F_b v_b - F_1\big[{A_1}^{-1}(A_b v_b + A_2 v_2)\big] + F_2 v_2,$$

y la restricción en (12) se convierte en

$$\begin{aligned} 0 =& \left([(F_b - F_1{A_1}^{-1}A_b)v_b - \widehat{\delta}]'\Sigma^{-1}[(F_b - F_1{A_1}^{-1}A_b)v_b - \widehat{\delta}] - c^{1-\alpha}\right) \\ &+ 2\left([(F_b - F_1{A_1}^{-1}A_b)v_b - \widehat{\delta}]'\Sigma^{-1}[(F_2 - F_1{A_1}^{-1}A_2)\right)v_2 \\ &+ v_2'\left([(F_2 - F_1{A_1}^{-1}A_2)]'\Sigma^{-1}[(F_2 - F_1{A_1}^{-1}A_2)]\right)v_2. \end{aligned}$$

Se trata de una expresión cuadrática para (el escalar) v_2 en términos de v_b y, definiendo los escalares d_0, d_1 y d_2 como,

$$\begin{aligned} d_0 &= \big[(F_2 - F_1{A_1}^{-1}A_2)\big]'\Sigma^{-1}\big[(F_2 - F_1{A_1}^{-1}A_2)\big], \\ d_1(v_b) &= 2\left([(F_b - F_1{A_1}^{-1}A_b)v_b - \widehat{\delta}]'\Sigma^{-1}[(F_2 - F_1{A_1}^{-1}A_2)\right), \text{and} \\ d_2(v_b) &= \left([(F_b - F_1{A_1}^{-1}A_b)v_b - \widehat{\delta}]'\Sigma^{-1}[(F_b - F_1{A_1}^{-1}A_b)v_b - \widehat{\delta}] - c^{1-\alpha}\right) \end{aligned}$$

se simplifica a $d_0 v_2^2 + d_1(v_b)v_2 + d_2(v_b) = 0$.

Utilizando la fórmula cuadrática, es posible resolver v_2 resolviendo el problema de minimización,

$$v_2(v_b) = \frac{-d_1(v_b) \pm \sqrt{d_1(v_b)^2 - 4d_0 d_2(v_b)}}{2d_0}. \tag{15}$$

Nótese que, por definición, $v_2 = v_{p^*}$.

Además si, $p^* = |\mathcal{N}| + 1$, v_b está vacío y, por tanto, v_2 no depende de v_b. (15) da lugar a dos soluciones, v_2^+ y v_1^-, correspondientes a la ambigüedad en los signos en (15). Se elige la solución v_2^* con el valor absoluto más pequeño.

Si $p^* > |\mathcal{N}| + 1$, , la optimización restringida en (11)-(13) es equivalente a la siguiente:

$$v_2^2 = \min_{v_b} \min_{\{+,-\}} \left(\frac{-d_1(v_b) \pm \sqrt{d_1(v_b)^2 - 4d_0 d_2(v_b)}}{2d_0} \right)^2, \tag{16}$$

donde la minimización interna es sobre el signo en la fórmula cuadrática.

En este punto, se tiene tanto v_2^* como v_b^*. Recuperando v_1^* mediante (14), se obtiene $v^* = [v_b^*, v_1^*, v_2^*]$.

E. Imputación de variables de política

La estimación de estudios de eventos en diseños de panel requiere establecer supuestos sobre el comportamiento de la variable de política fuera del intervalo de tiempo observado. En la sección 3.1 de este documento, se estimaron los estudios de evento en diseño de panel suponiendo que no hay cambios no observados en la variable de política fuera del periodo de estimación. Este esquema de imputación se implementa utilizando la opción `impute(nuchange)`.

`xtevent` permite otros esquemas para imputar la variable de política. Por ejemplo, `xtevent` puede suponer que la variable de política sigue una adopción escalonada, utilizando la opción `impute(stag)`. También puede imputar valores faltantes de la variable de política dentro del intervalo de fechas observado utilizando la opción `impute(instag)`.

Para ilustrar estas opciones, se utiliza un ejemplo de datos simulados de la sección 3 del documento y se muestran las variables dummy implícitas del tiempo de evento con los distintos esquemas de imputación. Para el ejemplo, se añaden algunos valores faltantes a la unidad 19. A continuación, se diferencia la variable de política. `xtevent` utiliza los rezagos y adelantos de la variable de política diferenciada para generar las variables dummy de tiempo de evento, siguiendo la ecuación (2).

En primer lugar, hay que pedirle a `xtevent` que genere las variables dummy de tiempo de evento sin ninguna imputación y se especifica la opción `savek(stub, noestimate)`) para guardarlas sin estimar el modelo.

```
. use simulation_data_dynamic.dta, clear
. qui xtset id t
. qui replace z=. if id==19 & (t==35 | t>=39)
. qui gen z_d=d.z
. qui xtevent y x, panelvar(id) timevar(t) policyvar(z)
>  window(5) savek(v, noestimate)
```

Las variables dummy de tiempo de evento con un prefijo "v" y que terminan en `m#` o `p#` corresponden a adelantos y rezagos de la variable de política diferenciada, como se describe en la sección A de este documento. El cuadro que aparece abajo muestra estas variables dummy de tiempo de evento para la unidad 19 en algunos períodos.

```
. list id t z z_d v_eq_m6 -v_eq_m1 if id==19 & t>=29, ///
> separator(4) noobs
```

id	t	z	z_d	v_eq_m6	v_eq_m5	v_eq_m4	v_eq_m3	v_eq_m2	v_eq_m1
19	29	0	0	1	0	0	0	0	0
19	30	0	0	.	.	0	0	0	0
19	31	0	0	1	.	.	0	0	0
19	32	0	0	0	1	.	.	0	0
19	33	0	0	0	0	1	.	.	0
19	34	0	0	.	.	0	1	.	.
19	35	.	.	.	.	.	0	1	.
19	36	0	.	.	.	.	.	0	1
19	37	1	1	.	.	.	.	.	0
19	38	1	0	.	.	.	.	.	.
19	39	.	.	.	.	.	.	.	.
19	40	.	.	.	.	.	.	.	.

Obsérvese que las variables dummy de tiempo de evento tienen valores faltantes en la parte inferior del cuadro porque no se ha hecho ningún supuesto sobre la variable de política fuera del intervalo de tiempo observado. Además, nótese que las variables dummy de tiempo de evento tienen algunos valores faltantes dentro del intervalo de tiempo observado debido al valor faltante en la variable de política en el periodo 35. A partir de la ecuación (2), este último valor faltante se traduce en dos valores internos faltantes en las variables dummy de eventos y un valor faltante en el caso del extremo izquierdo.

Para imputar la variable de política bajo la adopción escalonada, se utiliza la opción, `impute(stag)`. `xtevent` verifica que la variable de política sigue la adopción escalonada. Si es así, `xtevent` imputes the policy variable outside the observed time range.imputa la variable de política fuera del intervalo de tiempo observado. A continuación, utiliza la variable de política imputada para generar las variables dummy ficticias de evento-tiempode tiempo de evento y los puntos finales. Se añade la subopción `saveimp`para guardar la variable de política imputada como `z_imputed`. También se diferencia la nueva variable de política imputada para ver cómo se traducen sus adelantos y rezagos en nuevas variables dummy de tiempo de evento.

```
. cap drop v*

. qui xtevent y x, panelvar(id) timevar(t) policyvar(z) ///
> window(5) savek(v, noestimate) impute(stag, saveimp)

. qui gen z_imputed_d=d.z_imputed
```

A continuación, se compara la variable de política original, la variable de política imputada, la variable de política imputada diferenciada, algunas variables dummy de tiempo de evento y el punto final izquierdo generado utilizando la variable de política imputada. En primer lugar, la variable de política se ha imputado en el intervalo de tiempo observado. No obstante, la imputación también supone que la variable de política en los periodos posteriores a $t = 40$ tendría el mismo valor que la de ese último periodo. Esta imputación puede observarse en las variables dummy de tiempo de evento, que ahora tienen ceros que corresponden a las derivadas de la variable de política diferenciada en los periodos posteriores a 40.

```
. list id t z z_imputed z_imputed_d v_eq_m6 -v_eq_m3 if id==19 & t>=29, ///
> separator(4) noobs ab(11)
```

id	t	z	z_imputed	z_imputed_d	v_eq_m6	v_eq_m5	v_eq_m4	v_eq_m3
19	29	0	0	0	1	0	0	0
19	30	0	0	0	.	.	0	0
19	31	0	0	0	1	.	.	0
19	32	0	0	0	0	1	.	.
19	33	0	0	0	0	0	1	.
19	34	0	0	0	0	0	0	1
19	35	.	.	.	0	0	0	0
19	36	0	0	.	0	0	0	0
19	37	1	1	1	0	0	0	0
19	38	1	1	0	0	0	0	0
19	39	.	1	0	0	0	0	0
19	40	.	1	0	0	0	0	0

Ahora se pide `xtevent` que impute la variable de política utilizando la opción `impute(instag)`. Este esquema de imputación permite imputar los valores que faltan en la variable de política fuera y dentro del intervalo de tiempo observado. Como se describe en la sección A de este documento, la opción `impute(instag)` implementa la opción `impute(stag)` pero también imputa los valores faltantes dentro del intervalo de tiempo observado en los casos en los que es posible suponer algún valor basándose en los valores de la política en los periodos circundantes. Como en el ejemplo anterior, también se generó la variable de política imputada diferenciada para su comparación.

```
. cap drop v* z_imputed z_imputed_d

. qui xtevent y x, panelvar(id) timevar(t) policyvar(z) ///
> window(5) savek(v, noestimate) impute(instag, saveimp)

. qui gen z_imputed_d=d.z_imputed
```

A continuación, se compara la variable de política original, la variable de política imputada, la variable de política imputada diferenciada, algunas variables dummy de tiempo de evento y el punto final izquierdo generado con la variable de política imputada. En primer lugar, la variable de política imputada no tiene valores faltantes dentro o fuera del intervalo tiempo de evento. Al igual que en el ejemplo que utiliza `impute(stag)`, las variables dummy de tiempo de evento tienen ceros que corresponden a los adelantos de la variable de política diferenciada en períodos superiores a 40. Además, ahora las variables dummy de tiempo de evento no tienen valores faltantes dentro del intervalo de tiempo de evento.

```
. list id t z z_imputed z_imputed_d v_eq_m6 -v_eq_m3 if id==19 & t>=29, ///
> separator(4) noobs ab(11)
```

id	t	z	z_imputed	z_imputed_d	v_eq_m6	v_eq_m5	v_eq_m4	v_eq_m3
19	29	0	0	0	1	0	0	0
19	30	0	0	0	1	0	0	0
19	31	0	0	0	1	0	0	0
19	32	0	0	0	0	1	0	0
19	33	0	0	0	0	0	1	0
19	34	0	0	0	0	0	0	1
19	35	.	0	0	0	0	0	0
19	36	0	0	0	0	0	0	0
19	37	1	1	1	0	0	0	0
19	38	1	1	0	0	0	0	0
19	39	.	1	0	0	0	0	0

19	40	.	1	0	0	0	0	0

F. Estimación en bases de datos de corte transversal repetido

`xtevent` permite la estimación con bases de datos de corte transversal repetido cuando `policyvar` varía a nivel de grupo, y `panelvar` identifica los grupos. Por ejemplo, `panelvar` podría indicar los estados en los que cambia `policyvar` mientras que las observaciones en la base de datos podrían ser individuos en cada estado. `xtevent` permite estimaciones en estas configuraciones directamente con la opción. `repeatedcs`. También permite utilizar el procedimiento de dos pasos descrito en Hansen (2007). Para utilizar este último método, el usuario debe usar primero el comando `get_unit_time_effects` para estimar los efectos unidad-tiempo y luego utilizar estas estimaciones como entrada para `xtevent`.

Se ilustra el uso de `get_unit_time_effects`. En primer lugar, se crea una variable `state` que representa los grupos en los que los individuos reciben el tratamiento en el mismo periodo. A continuación, se llama `get_unit_time_effects`. Guarda un archivo `dta` con los efectos de tiempo unitario.

```
. gen state=eventtime

. xtset, clear

. get_unit_time_effects y x, panelvar(state) timevar(t)
> saving("effect_file.dta", replace)
  (output omitted)
```

Luego, se mantiene una observación por estado-tiempo en los datos de corte transversal repetido y se fusiona la base de datos con los efectos de tiempo unitario. Después, se ejecuta `xtevent`. Dado que se usa una base de datos más pequeño para estimar el evento-estudio, este método puede ser más rápido que utilizar la opción `repeatedcs`.

```
. qui bysort state t (z): keep if _n==1

. keep state t z

. qui merge m:1 state t using effect_file.dta, nogen

. xtevent _unittimeeffects, panelvar(state) timevar(t) policyvar(z) window(5)
  (output omitted)
```

G. Estimación con efectos de tratamiento heterogéneos en contextos de adopción escalonada

`xtevent` permite la estimación de efectos de tratamiento heterogéneos en entornos de adopción escalonada. Siguiendo el enfoque de Sun y Abraham (2021), introducimos un escenario con efectos de tratamiento heterogéneos entre cohortes de tratamiento de la siguiente manera: sea $t^\star(i)$ la cohorte de la unidad i, es decir, el periodo en el que la unidad i adopta la política. Denótense los efectos de la política sobre el resultado de la cohorte $t^\star$ as $\{\beta_{m,t^\star}\}_{m=-G}^{M}$. La ecuación para el resultado se convierte en:

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

Para estimar la trayectoria evento-estudio del resultado en este contexto, `xtevent` estima una versión ampliada de la ecuación (2) interactuando las variables dummy de tiempo de evento con indicadores de cohorte, como se propone en Freyaldenhoven et al. (2021):

$$y_{it} = \sum_{c} \sum_{k=-G-L_G}^{M+L_M-1} \left(\mathbf{1}\left\{t^\star(i) = c\right\}\right)\left(\delta_{k,c}\Delta z_{i,t-k} + \delta_{M+L_M,c} z_{i,t-M-L_M} + \delta_{-G-L_G-1,c}(1 - z_{i,t+G+L_G})\right) + \alpha_i + \gamma_t + q_{it}'\psi + C_{it} + \varepsilon_{it}, \tag{17}$$

Se utiliza el estimador de efectos fijos bidireccionales para estimar (17), que equivale al estimador propuesto en Sun y Abraham (2021). Además, `xtevent` también permite estimar (17)utilizando cualquiera de las estrategias de estimación descritas en la sección 2.1 excepto la estimación IV.

`xtevent` estima (17) en una muestra definida por las opciones de `cohort` y `control_cohort`. en una muestra definida por las opciones de `cohort(create)` pide a `xtevent` que genere automáticamente una variable categórica para las cohortes de tratamiento. `xtevent` establece el valor de esta variable categórica al valor de la variable de tiempo la primera vez que se trata una unidad. Para las unidades que nunca son tratadas, `xtevent` sestablece el valor de esta variable como faltante. De forma similar, la opción `control_cohort(create)` pide a `xtevent` tque genere automáticamente un indicador para la cohorte de control. `xtevent` establece este indicador a 1 para las unidades nunca

tratadas y a cero en caso contrario.

Si la cohorte en `control_cohort` ies una cohorte nunca tratada, `xtevent` estima la ecuación (17) en toda la muestra. En caso contrario, `xtevent` estima (17) en el subconjunto de periodos en los que las observaciones de la cohorte de control aún no han sido tratadas. Por defecto, la estimación excluye las cohortes siempre tratadas.

Tras obtener las estimaciones $\hat{\delta}_{k,c}$ para $k = -G - L_G - 1, ..., M + L_M$ y para cada cohorte c, `xtevent` obtiene la estimación del efecto medio del tratamiento en el tiempo de evento k, $\hat{\delta}_k$, como media de las estimaciones $\hat{\delta}_{k,c}$ponderando por la proporción de observaciones de la cohorte c en cada tiempo relativo k (Sun y Abraham, 2021):

$$\hat{\delta}_k = \sum_c \delta_{k,c} \hat{Pr} \left\{ t^\star(i) = c \mid t^\star(i) \in [-k, T - k] \right\}.$$

La varianza de $\hat{\delta}_k$ se obtiene utilizando las fórmulas del Apéndice C.1 de Sun y Abraham (2021).

Para permitir una mayor flexibilidad `xtevent` también permite la estimación de (17) fuera de entornos de adopción escalonada o en entornos en los que el usuario desea agregar cohortes de tratamiento para variables de cohorte y cohorte de control proporcionadas por el usuario a través de las opciones `cohort(variable varname, [,force])` y `control_cohort(variable varname, [,force])`.