



GRADO EN CIENCIA DE DATOS



VNIVERSITAT  
DE VALÈNCIA

TRABAJO FIN DE GRADO

---

MÉTODOS DE ANÁLISIS DE EFECTIVIDAD DE  
MEDIDAS: EVALUACIÓN DEL IMPACTO DEL USO  
DE MASCARILLAS MEDIANTE DIFERENCIA EN  
DIFERENCIAS Y SYNTHETIC CONTROL

---

AUTOR: ENRIQUE SAYAS BAILACH

TUTOR: CARLES BRETÓ

JULIO 2025





VNIVERSITAT  
DE VALÈNCIA



Escola Tècnica Superior  
d'Enginyeria **ETSE-UV**

## TRABAJO FIN DE GRADO

---

# MÉTODOS DE ANÁLISIS DE EFECTIVIDAD DE MEDIDAS: EVALUACIÓN DEL IMPACTO DEL USO DE MASCARILLAS MEDIANTE DIFERENCIA EN DIFERENCIAS Y SYNTHETIC CONTROL

---

**AUTOR: ENRIQUE SAYAS BAILACH**

**TUTOR: CARLES BRETÓ**

---



### **Declaración de autoría:**

Yo, Enrique Sayas Bailach, declaro la autoría del Trabajo Fin de Grado titulado “Métodos de análisis de efectividad de medidas: evaluación del impacto del uso de mascarillas mediante diferencia en diferencias y synthetic control” y que el citado trabajo no infringe las leyes en vigor sobre propiedad intelectual. El material no original que figura en este trabajo ha sido atribuido a sus legítimos autores.

Valencia, 25 de junio de 2025

A handwritten signature in black ink that reads "Enrique S.B.". The signature is written in a cursive style, with the first letter 'E' being large and stylized. The entire signature is enclosed within a hand-drawn, irregular oval loop.

Fdo: Enrique Sayas Bailach



---

**Resumen:**

El uso de mascarillas ha sido una medida fundamental para el control de la transmisión de COVID-19. En este contexto, la investigadora Tori L. Cowger lideró en *The New England Journal of Medicine* un estudio que evaluaba el impacto del levantamiento del mandato de uso obligatorio de mascarillas en centros escolares del estado de Massachusetts. En concreto, se compararon los distritos que retiraron la obligatoriedad del uso de mascarillas en las semanas posteriores a la rescisión del mandato estatal con aquellos que decidieron mantenerla. Sin embargo, el código relativo al artículo y a los resultados obtenidos no ha sido publicado. Esta limitación condiciona la capacidad de extender el estudio original para conocer el impacto real del uso de mascarillas, así como la elaboración de nuevos estudios que ayuden a cuantificar este impacto. De este modo, este trabajo tiene como objetivo principal reproducir la metodología del estudio original a partir de la documentación complementaria facilitada por la autora, con el objetivo último de contribuir a la comprensión y evaluación del impacto del uso de mascarillas en el entorno escolar. Para ello, se desarrolla una implementación propia del análisis de diferencia en diferencias. Además, se comparan los resultados obtenidos con métodos alternativos, como el enfoque de Generalized Synthetic Control, con el fin de valorar la robustez de las conclusiones originales y explorar posibles mejoras metodológicas.

---





---

**Abstract:**

The use of face masks has been a fundamental measure in controlling the transmission of COVID-19. In this context, researcher Tori L. Cowger led a study published in *The New England Journal of Medicine* that evaluated the impact of lifting the mandatory mask mandate in schools across the state of Massachusetts. Specifically, the study compared school districts that removed the mask requirement in the weeks following the repeal of the statewide mandate with those that chose to maintain it. However, the code related to the article and the resulting estimates has not been made publicly available. This limitation affects the ability to extend the original study in order to better understand the actual impact of mask use, as well as to develop new studies that help quantify this effect. Accordingly, the main objective of this work is to reproduce the methodology of the original study based on the supplementary documentation provided by the author, with the ultimate aim of contributing to the understanding and evaluation of the impact of mask use in school settings. To this end, a custom implementation of the difference-in-differences analysis is developed. In addition, the results are compared with alternative methods, such as the Generalized Synthetic Control approach, in order to assess the robustness of the original conclusions and explore potential methodological improvements.

---



---

**Resum:**

L'ús de les mascaretes ha sigut una mesura fonamental per al control de la transmissió de la COVID-19. En aquest context, la investigadora Tori L. Cowger va liderar un estudi publicat en *The New England Journal of Medicine* en què s'avaluava l'impacte de l'eliminació del mandat d'ús obligatori de mascareta en els centres escolars de l'estat de Massachusetts. En concret, es compararen els districtes que retiraren l'obligatorietat de la mascareta en les setmanes posteriors a la rescissió del mandat estatal amb aquells que decidiren mantindre-la. Tanmateix, el codi relacionat amb l'article i amb els resultats obtinguts no ha sigut publicat. Aquesta limitació condiciona la capacitat d'estendre l'estudi original per tal de conèixer l'impacte real de l'ús de mascaretes, així com l'elaboració de nous estudis que permeten quantificar aquest impacte. Així, aquest treball té com a objectiu principal reproduir la metodologia de l'estudi original a partir de la documentació complementària facilitada per l'autora, amb l'objectiu últim de contribuir a la comprensió i avaluació de l'impacte de l'ús de mascaretes en l'àmbit escolar. Per a això, es desenvolupa una implementació pròpia de l'anàlisi de diferències en diferències. A més, es comparen els resultats obtinguts amb mètodes alternatius, com ara l'enfocament de Generalized Synthetic Control, amb la finalitat de valorar la robustesa de les conclusions originals i explorar possibles millores metodològiques.

---



---

**Agradecimientos:**

A mi tutor, Carles Bretó, por su labor y ayuda. Al resto de profesores del Grado de Ciencia de Datos. Y a mi familia por su apoyo, paciencia y fe.

---



# Índice general

<b>1. Introducción</b>	<b>17</b>
1.1. Introducción . . . . .	17
1.2. Objetivos . . . . .	18
<b>2. Datos</b>	<b>19</b>
2.1. Fuentes de datos . . . . .	19
2.1.1. Datos de incidencia de COVID-19 . . . . .	19
2.1.2. Datos descriptivos de los distritos . . . . .	20
2.2. Preprocesado de datos . . . . .	20
2.2.1. Lectura de los datos . . . . .	20
2.2.2. Corrección . . . . .	21
<b>3. Metodología</b>	<b>23</b>
3.1. Diseño de investigación . . . . .	23
3.1.1. Criterios de inclusión y exclusión de distritos . . . . .	23
3.1.2. Periodo de estudio y unidad de análisis . . . . .	24
3.1.2.1. Periodo de estudio . . . . .	24
3.1.2.2. Unidad de análisis . . . . .	25
3.2. Reproducción del enfoque original . . . . .	26
3.2.1. Diferencia en Diferencias . . . . .	26
3.2.1.1. Tratamiento escalonado . . . . .	27
3.2.1.2. Modelo de referencia . . . . .	29
3.2.1.3. Estimación del efecto . . . . .	29
3.2.1.4. Cumplimiento del supuesto de tendencias paralelas . . . . .	30
3.2.1.5. Implementación en R . . . . .	31
3.2.2. Estimaciones semanales y acumuladas del ATT . . . . .	32
3.3. Extensión metodológica . . . . .	33
3.3.1. Generalized Synthetic Control . . . . .	33
3.3.1.1. Modelo de referencia . . . . .	34

---

3.3.1.2.	Implementación en R . . . . .	34
3.3.2.	Estimaciones semanales y acumuladas del ATT . . . . .	34
3.3.3.	Comparación metodológica: ventajas, limitaciones y supuestos . . .	35
3.3.3.1.	Supuesto de tendencias paralelas . . . . .	36
3.3.3.2.	Flexibilidad temporal . . . . .	36
3.3.3.3.	Ventajas prácticas . . . . .	36
3.3.3.4.	Limitaciones . . . . .	36
<b>4.</b>	<b>Resultados</b>	<b>37</b>
4.1.	Resultados . . . . .	37
4.1.1.	Verificación del supuesto de tendencias paralelas . . . . .	37
4.1.2.	Resultados mediante Diferencia en Diferencias . . . . .	38
4.1.3.	Resultados mediante Generalized Synthetic Control . . . . .	39
4.2.	Discusión . . . . .	39
<b>5.</b>	<b>Conclusiones</b>	<b>41</b>
5.1.	Conclusiones . . . . .	41
5.2.	Trabajo futuro . . . . .	42
<b>A.</b>	<b>Anexos</b>	<b>43</b>
A.1.	Código . . . . .	43
	<b>Bibliografía</b>	<b>46</b>



# Capítulo 1

## Introducción

### 1.1. Introducción

Desde el inicio de la pandemia de COVID-19, las mascarillas se han convertido en una herramienta clave para reducir la transmisión del virus (Talic et al., 2021; Andrejko et al., 2022), especialmente en entornos cerrados y de alta interacción como los centros escolares (Gettings et al., 2021). La implementación de mandatos de uso obligatorio de mascarillas ha sido una de las medidas no farmacológicas más extendidas a nivel global (Masks4All, 2020), y su posterior retirada ha generado un intenso debate tanto en la comunidad científica como en el ámbito educativo y político (Branch-Elliman et al., 2022).

En este contexto, el estado de Massachusetts (Estados Unidos) constituye un caso de especial interés. Tras la rescisión del mandato estatal de mascarillas en febrero de 2022, los distritos escolares del estado dispusieron de autonomía para decidir si mantener o eliminar dicha medida (Decker, 2021; Boston Herald, 2022). Esta variabilidad dio lugar a un entorno natural para la evaluación de impacto de políticas sanitarias mediante técnicas de inferencia causal.

Sobre esta base, la investigadora Tori L. Cowger y su equipo publicaron en *The New England Journal of Medicine* un estudio que evaluaba el impacto del levantamiento del uso obligatorio de mascarillas en los centros escolares de Massachusetts (Cowger et al., 2022a). El análisis comparaba la evolución de casos de COVID-19 entre distritos que retiraron la medida y aquellos que la mantuvieron, aprovechando la adopción escalonada del tratamiento. El trabajo utilizó un modelo de diferencia en diferencias con grupo de control dinámico (Callaway and Sant'Anna, 2021b), mostrando un aumento significativo en la incidencia de casos en los distritos que retiraron la obligatoriedad.

Sin embargo, ni el código fuente ni los datos procesados del estudio original han sido publicados, lo que limita la posibilidad de verificar los resultados, reproducir el análisis y desarrollar extensiones metodológicas sobre el mismo. Esta carencia impide tanto la transparencia como la acumulación de evidencia en un tema con importantes implicaciones para la salud pública.

Con el objetivo principal de reproducir la metodología del estudio Cowger et al. (2022a) a partir de la documentación complementaria facilitada por los autores (Cowger et al., 2022c,b), en este trabajo se implementa un modelo propio de diferencia en diferencias con tratamiento escalonado (Callaway and Sant'Anna, 2021b), y se introduce además un segundo enfoque metodológico alternativo, el *generalized synthetic control* (Xu and Liu,

2021), con el fin de comparar los resultados y evaluar la robustez de las conclusiones bajo diferentes supuestos estadísticos (Baker et al., 2022).

Los resultados obtenidos con el modelo de diferencia en diferencias muestran un exceso estimado de 14.102 casos de COVID-19 atribuibles al levantamiento de las mascarillas durante las 15 semanas posteriores a la rescisión del mandato estatal, lo que representa un incremento de 53,5 casos por cada 1000 personas en los distritos tratados. En cambio, el método *generalized synthetic control* estima un exceso más moderado de 4.107 casos (15,6 por 1000), concentrado principalmente en el personal escolar. En el caso del alumnado, los efectos estimados por *generalized synthetic control* no alcanzan significación estadística.

Estas diferencias reflejan la sensibilidad de los resultados a los supuestos metodológicos de cada modelo. No obstante, ambos enfoques coinciden en señalar un aumento de la incidencia de casos tras la retirada de las mascarillas, especialmente en el personal docente y no docente. Por tanto, el trabajo no sólo valida parcialmente las conclusiones del estudio original, sino que también aporta un análisis comparado que contribuye a una comprensión más matizada del impacto real de esta medida sanitaria en el ámbito escolar.

## 1.2. Objetivos

El objetivo principal de este trabajo es reproducir la metodología del estudio original liderado por Tori L. Cowger (Cowger et al., 2022a), publicado en *The New England Journal of Medicine*, a partir de la documentación complementaria proporcionada por los autores (Cowger et al., 2022c,b). En particular, se busca desarrollar una implementación propia del modelo de diferencia en diferencias con tratamiento escalonado (Callaway and Sant'Anna, 2021b), aplicado al contexto del levantamiento de los mandatos de mascarilla en centros escolares del estado de Massachusetts (Boston Herald, 2022).

Como objetivos intermedios, se plantea la reproducción de los principales resultados del estudio original, incluyendo tablas y gráficos clave, así como la incorporación de un enfoque metodológico alternativo: el modelo *generalized synthetic control* (Xu and Liu, 2021). Este segundo modelo permite contrastar los resultados obtenidos bajo supuestos distintos, lo que refuerza el análisis de robustez metodológica.

En última instancia, el propósito final del trabajo es contribuir a la comprensión y evaluación del impacto del uso de mascarillas en el entorno escolar. A través de la comparación de enfoques y la discusión de sus implicaciones, se pretende aportar evidencia adicional sobre el efecto de esta medida de salud pública y promover la utilización de técnicas reproducibles y transparentes en la evaluación de políticas sanitarias.

El resto del trabajo se estructura de la siguiente manera: en la Sección 2 se describen las fuentes de datos utilizadas y el proceso de preprocesado aplicado; en la Sección 3 se detalla la metodología seguida para la implementación de los modelos; en la Sección 4 se presentan los resultados y su comparación; y finalmente, en la Sección 5 se exponen las conclusiones y posibles líneas de trabajo futuro.

# Capítulo 2

## Datos

### 2.1. Fuentes de datos

#### 2.1.1. Datos de incidencia de COVID-19

Como se indica en Cowger et al. (2022a), los datos de incidencia de COVID-19 han sido extraídos de los informes publicados por el Massachusetts Department of Elementary and Secondary Education (2023). La inspección inicial de dichos datos muestra la existencia de múltiples factores que aconsejan extremar el cuidado a la hora de su análisis. Entre estos factores destacan algunos problemas en fechas concretas, incluyendo datos faltantes que hacen que la frecuencia sea irregular y algunos probables errores de medida. También cabe tener en cuenta la representatividad del período para el que se disponen datos, así como la pertinencia de las variables para las que se muestran observaciones en los informes, los cuales mayoritariamente mantienen un mismo formato durante todo el período.

El formato de los informes es en PDF, con una estructura semitabular distribuida en varias páginas con una fila por distrito escolar. Para cada distrito y semana de reporte, se incluyen las siguientes variables:

- Código de identificación del distrito escolar.
- Nombre del distrito escolar.
- Número de casos positivos confirmados entre estudiantes.
- Número de casos positivos confirmados entre el personal escolar (docente y no docente).
- Número total de pruebas diagnósticas realizadas.
- Número de pruebas con resultado positivo.
- Tasa de positividad de las pruebas diagnósticas (%).

Los primeros dos informes disponibles (correspondientes a las semanas iniciales del curso) no incluyen información sobre el número de pruebas realizadas.

La frecuencia de publicación de los informes era semanal, con los datos recopilados cada miércoles a las 17:00h, y referidos al periodo jueves-miércoles de cada semana epidemiológica (Cowger et al., 2022c).

Las semanas de estudio incluidas son un total de 40, que abarcan desde principios de septiembre de 2021 hasta mediados de junio de 2022 (Boston Public Schools, 2021). Este periodo cubre tanto la etapa en la que el uso de mascarillas era obligatorio en todos los distritos escolares, como la fase posterior al 28 de febrero, cuando la normativa se flexibilizó y los distritos comenzaron a levantar esta medida de forma escalonada (Boston Herald, 2022).

En los registros con 0 casos, tanto en estudiantes como en personal, Massachusetts Department of Elementary and Secondary Education (2023) no distingue entre la ausencia real de contagios y la falta de envío de datos por parte del distrito escolar correspondiente (Cowger et al., 2022c). Esta ambigüedad dificulta la interpretación directa de los ceros, ya que no es posible saber si representan un valor legítimo o simplemente un reporte ausente.

Esta limitación adquiere especial relevancia en el análisis de impacto, dado que una distribución desigual de registros sin notificación entre los grupos tratados y no tratados podría afectar la estimación de los efectos del levantamiento del uso obligatorio de mascarillas. Por ejemplo, si los distritos que retiraron la medida informaron con menor regularidad, los resultados podrían subestimar el efecto real del tratamiento.

### 2.1.2. Datos descriptivos de los distritos

Además de los datos de incidencia de COVID-19, se recopilaron datos descriptivos que caracterizan a los distritos escolares desde el punto de vista demográfico, socioeconómico y estructural.

Los datos provienen de Massachusetts Department of Elementary and Secondary Education (2024), que proporciona estadísticas anuales de matrícula y características del alumnado y del personal para cada distrito escolar.

Las principales variables consideradas en esta categoría incluyen:

- Matrícula total de estudiantes.
- Número de empleados del distrito (docentes y no docentes).
- Porcentaje de estudiantes con características específicas:
  - Estudiantes de bajos ingresos.
  - Estudiantes con necesidades educativas especiales.
  - Estudiantes aprendices del idioma inglés.
- Composición étnico-racial del alumnado y del personal:
  - Porcentaje de estudiantes y personal de origen afroamericano, latino, asiático, blanco no hispano, y otros grupos.

## 2.2. Preprocesado de datos

### 2.2.1. Lectura de los datos

Para la lectura de los datos se ha llevado a cabo una función cuya lógica aprovecha la estructura de los informes semanales publicados por el Massachusetts Department of

Elementary and Secondary Education (2023).

La función (A.1) recorre una lista con las rutas a los archivos PDF, extrae la fecha de cada uno de ellos, y convierte el contenido tabular en un data frame estructurado. El resultado es un único conjunto de datos que consolida la información semanal por distrito escolar.

A continuación, se muestra la lógica general de la función:

1. Extracción de fechas: Se utiliza una expresión regular para identificar la fecha del informe directamente desde el nombre del archivo. Esta fecha se asignará posteriormente a cada registro como variable Date.
2. Inicialización del data frame: Se crea una estructura vacía con las columnas deseadas: código del distrito, nombre, número de casos en estudiantes y personal, número total de pruebas PT, pruebas positivas PPT y ratio de positividad PPR.
3. Lectura del contenido del PDF: Se aplica la función `pdf_text()` del paquete `pdftools` (Ooms, 2024) para convertir el PDF a texto legible, separado por páginas.
4. Extracción de los datos por página: Se recorre el contenido de cada página a partir de la segunda, ya que es donde comienzan los datos relevantes. Las líneas de texto se dividen por saltos de línea y posteriormente se separan en columnas utilizando múltiples espacios.
5. Filtrado y limpieza: Se eliminan las filas vacías o incompletas (sin código de distrito) y se asigna la fecha del informe a cada fila.
6. Detección del final del informe: El bucle de lectura se interrumpe al detectar la página que contiene la cabecera Education Collaboratives, que marca el final de los distritos escolares de interés para este estudio.
7. Consolidación del conjunto de datos: Todos los registros extraídos de las distintas páginas y archivos se concatenan en un único data frame.

### 2.2.2. Corrección

Las correcciones aplicadas a los datos están basadas en las indicaciones planteadas por Tori L. Cowger en Cowger et al. (2022c). En él se detalla cómo se debe proceder frente a los datos acumulados de dos semanas, los registros con 0 casos notificados y las semanas posteriores a la reincorporación del uso obligatorio de la mascarilla en el distrito de Brookline.

En relación con los datos acumulados de dos semanas (A.2), durante las semanas que incluyen vacaciones escolares, DESE optó por publicar un único informe que cubre el número de positivos para dos semanas.

En estos casos, se ha asumido un mismo ratio para ambas. Es decir, si un distrito reportó 100 casos en una semana acumulada, se imputan 50 casos a cada una de las dos semanas correspondientes.

Con respecto a los registros con 0 casos notificados (A.3), cuando un distrito no enviaba sus datos dentro del plazo estipulado, el sistema asignaba automáticamente un valor de cero casos tanto para estudiantes como para personal. Por contra, esta característica

limita la posibilidad de conocer si el no reporte se debe a que han habido 0 casos o a que directamente no han sido reportados.

Para identificar estos ceros imputados, se utilizaron los siguientes criterios:

- La suma entre casos en estudiantes y en personal es igual a 0 pero hay al menos un test positivo, o
- La suma entre casos en estudiantes y en personal es igual a 0 pero la media de las 4 semanas anteriores es mayor a 10.

Los registros en los que se consideraron ausencia de reporte se trataron de la siguiente forma:

1. Se elimina el registro del conjunto de datos, y
2. el ratio de la semana siguiente se calcula para la población de dos semanas.

Asimismo, el distrito de Brookline reincorporó la obligatoriedad del uso de las mascarillas el 23 de mayo de 2022 (Brookline Town News, 2022). Para este distrito escolar se excluyeron las semanas de estudio posteriores a esta fecha (las últimas 3 semanas del periodo de estudio).

# Capítulo 3

## Metodología

### 3.1. Diseño de investigación

#### 3.1.1. Criterios de inclusión y exclusión de distritos

Los criterios de inclusión y exclusión de distritos han sido extraídos de la información proporcionada en Cowger et al. (2022c). En él se detallan los criterios geográficos, institucionales y de calidad del dato a tener en cuenta para la correcta selección de distritos final. A continuación se reproduce y resume la descripción del procedimiento y pasos implementados en Cowger et al. (2022a) con el fin de tener una completa exposición de ideas en este documento.

Inicialmente, se tienen todos los distritos escolares públicos del estado de Massachusetts que reportaron datos al DESE durante el curso escolar 2021–2022.

En base a los criterios geográficos, se seleccionaron únicamente los distritos escolares ubicados dentro del área metropolitana de Boston, definida oficialmente como la división NECTA (New England City and Town Area) de Boston–Cambridge–Newton (Proximity-One, 2013).

Respecto a los criterios institucionales, se excluyeron del análisis los siguientes tipos de entidades escolares:

- Escuelas charter (autónomos).
- Escuelas vocacionales y técnicas.
- Collaboratives educativos y escuelas especiales.

Finalmente, a partir de los criterios de calidad del dato, se eliminaron aquellos distritos con inconsistencias en los casos registrados. En particular, se excluyeron:

- Distritos con más de 10 semanas con 0 casos reportados tanto en estudiantes como en personal.
- Distritos con al menos 5 semanas donde se registran 0 casos, pero a la vez presentan:
  - Resultados positivos en pruebas agrupadas (pooled testing), o bien

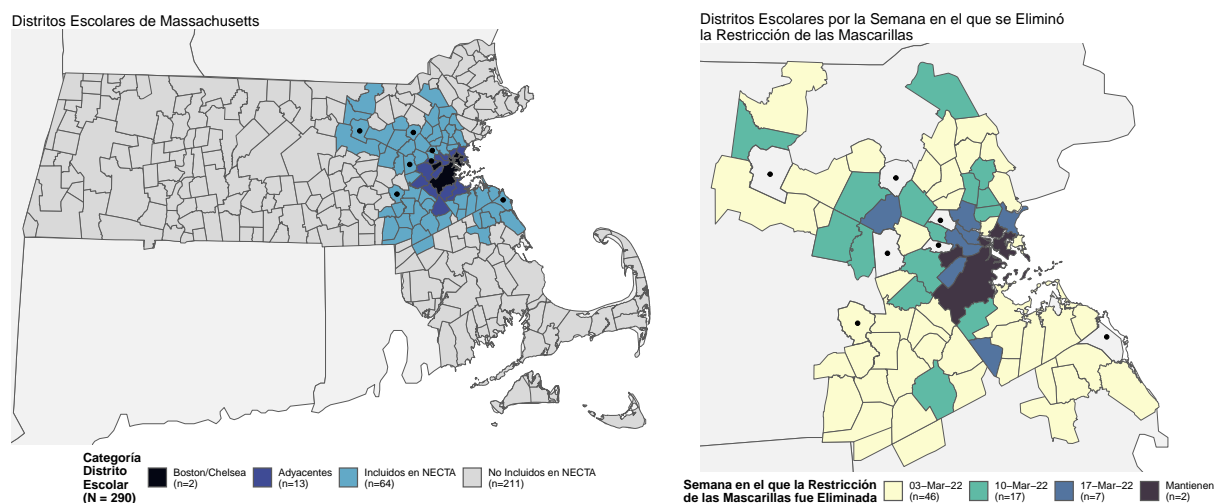
- Un promedio superior a 10 casos en las cuatro semanas previas.

De este modo, la selección de distritos final quedó conformada por 72 distritos escolares públicos tradicionales, que constituyen la unidad de análisis a lo largo de 40 semanas escolares.

A continuación, se muestra un resumen de los criterios de inclusión y exclusión aplicados a la muestra inicial de 400 distritos:

- Se seleccionaron los 290 distritos públicos tradicionales.
- Se restringió geográficamente a 79 distritos dentro de la división Boston–Cambridge–Newton (NECTA).
- Se eliminaron 7 distritos adicionales por mala calidad de los datos.

Figura 3.1: Selección de distritos escolares



Fuente: elaboración propia del original en Cowger et al. (2022a).

### 3.1.2. Periodo de estudio y unidad de análisis

#### 3.1.2.1. Periodo de estudio

El periodo de estudio comprende el curso escolar completo 2021–2022 en Massachusetts, que abarca un total de 40 semanas académicas consecutivas (Cowger et al., 2022a). El calendario escolar se extiende desde principios de septiembre de 2021 hasta mediados de junio de 2022 (Boston Public Schools, 2021).

Este periodo comprende semanas anteriores y posteriores al levantamiento de la obligatoriedad de las mascarillas:

- Fase pretratamiento: hasta el 28 de febrero de 2022, todos los distritos escolares mantenían la obligatoriedad de uso de mascarillas impuesta a nivel estatal.



- Posteriormente, entre marzo y abril de 2022, los distintos distritos levantaron la medida de forma escalonada (Boston Herald, 2022), lo que permite explotar la variabilidad temporal del tratamiento en un diseño de diferencias en diferencias con implementación escalonada.
- Finalmente, el periodo posterior (marzo–junio) permite observar los efectos acumulados a medio plazo del levantamiento de las medidas.

### 3.1.2.2. Unidad de análisis

Siguiendo la metodología de Cowger et al. (2022a), la unidad de análisis principal es la observación distrito escolar–semana ponderada por la población del distrito.

Los modelos tradicionales de diferencia en diferencias tratan a todas las observaciones (distritos escolares) por igual. De este modo, para obtener resultados más robustos, se busca encontrar el efecto a nivel poblacional.

A continuación se reproduce la Tabla S7 de Cowger et al. (2022c), que presenta el número de casos de COVID-19 entre estudiantes y personal escolar, desglosado según la fecha en que cada distrito eliminó la obligatoriedad del uso de mascarillas. Los datos se agrupan en tres periodos clave: pre-Ómicron (semanas 1–10), periodo Ómicron (semanas 11–25) y semanas posteriores al levantamiento del mandato estatal de mascarillas (semanas 26–40).

Tabla 3.1: Porcentaje de casos pre-Omicron, durante la ola Omicron BA.1, y después de la eliminación de la restricción de las mascarillas

Semana Eliminación Mascarilla	n Distritos	Pre-Omicron (Semanas 1-10, n= 10 semanas)	Omicron BA.1 (Semanas 11-25, n= 15 semanas)	Post-Levantamiento/ Omicron BA.2+ (Semanas 26-40, n=15 semanas)
Todos				
No levantaron	2	686	8343	4437
03-Mar-22	46	2750	33941	18089
10-Mar-22	17	1212	15142	13356
17-Mar-22	7	525	5276	4222
Todos	72	5173	62702	40104
Personal				
No levantaron	2	161	2519	1038
03-Mar-22	46	362	4896	3941
10-Mar-22	17	159	2227	2214
17-Mar-22	7	86	914	963
Todos	72	768	10556	8156
Estudiantes				
No levantaron	2	525	5824	3399
03-Mar-22	46	2388	29045	14148
10-Mar-22	17	1053	12915	11142
17-Mar-22	7	439	4362	3259
Todos	72	4405	52146	31948

Fuente: elaboración propia del original en Cowger et al. (2022c).

La reproducción de esta tabla a partir de los informes semanales en formato PDF publicados por el Massachusetts Department of Elementary and Secondary Education (2023) constata que el proceso de extracción y adecuación de los datos ha sido implementado correctamente. La correspondencia entre los resultados obtenidos y los valores presentados en el estudio original confirma la validez de la lógica de lectura y los criterios de clasificación aplicados, lo que refuerza la fiabilidad del conjunto de datos construido para el análisis.

## 3.2. Reproducción del enfoque original

### 3.2.1. Diferencia en Diferencias

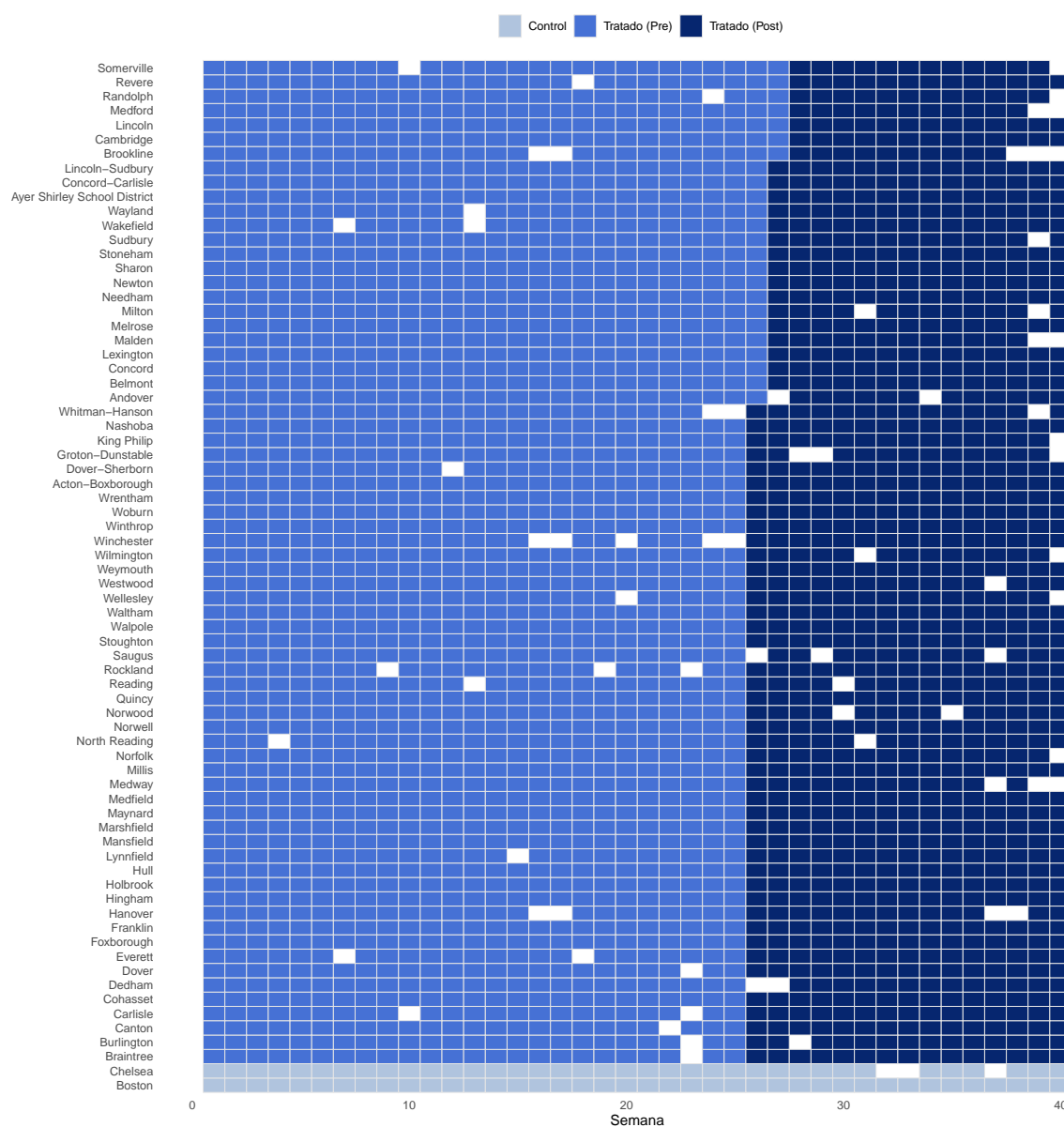
El método Diferencia en Diferencias (Callaway and Sant'Anna, 2021b) permite estimar el efecto causal de una intervención cuando no se aplica de forma simultánea en todas las unidades de análisis. En este caso, se parte de un contexto de tratamiento escalonado donde los distritos escolares levantaron los mandatos de mascarilla en distintas semanas (Boston Herald, 2022). A partir de su implementación en R (Callaway and Sant'Anna,

2021a), es posible obtener la estimación del efecto de mantener el uso de mascarillas en el ámbito escolar, así como verificar el cumplimiento del supuesto de tendencias paralelas para la validez de los resultados.

### 3.2.1.1. Tratamiento escalonado

El tratamiento escalonado se debe a que los distintos distritos escolares no aplicaron la retirada del uso obligatorio de mascarillas en el mismo momento (Boston Herald, 2022). Esta variación temporal genera la necesidad de un enfoque metodológico que tenga en cuenta las diferentes fechas de intervención para cada unidad de análisis.

Figura 3.2: Estado del tratamiento en los distritos estudiados



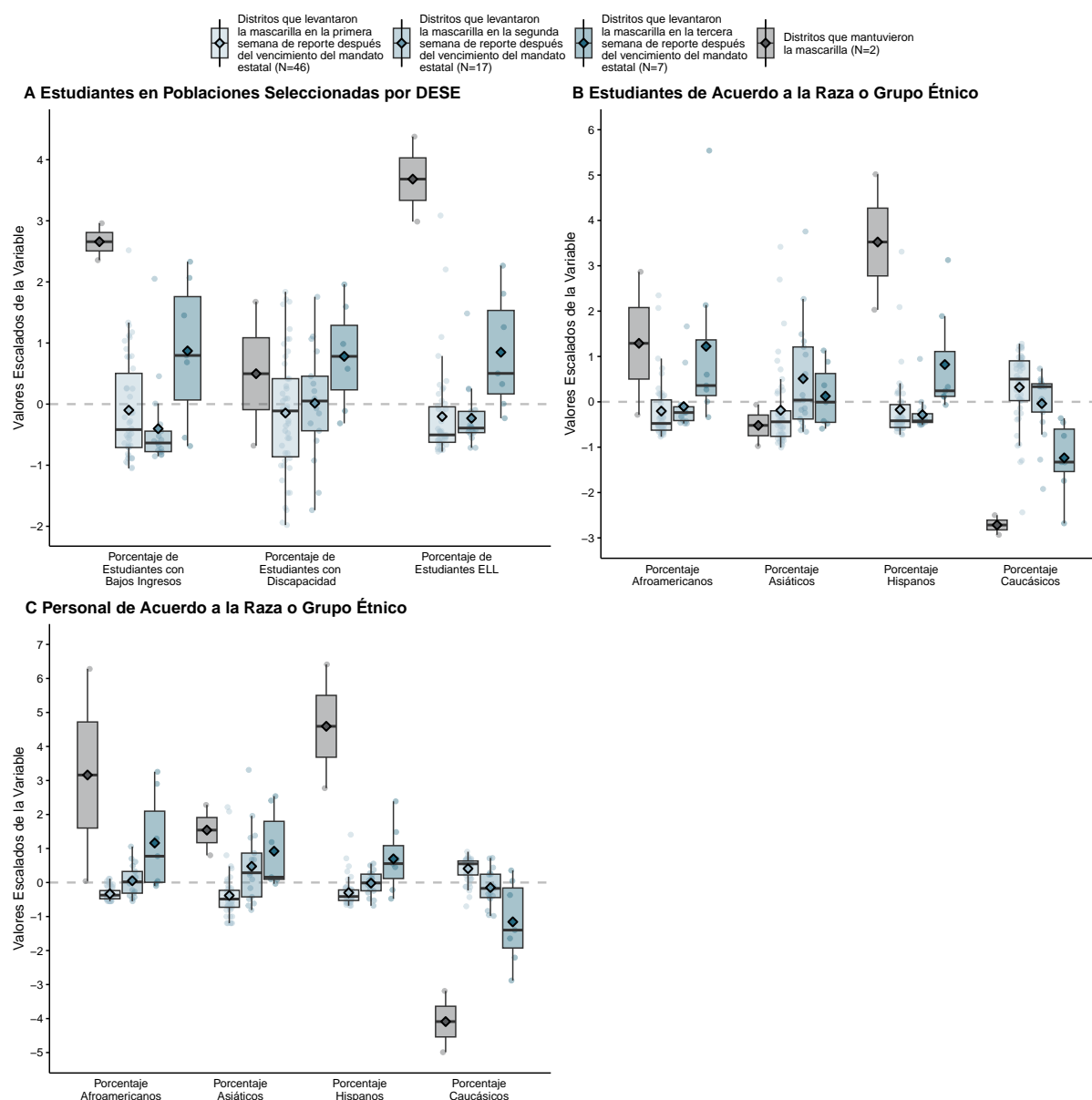
Fuente: elaboración propia.

Los distritos que mantuvieron durante más tiempo el uso obligatorio de mascarillas tendían a presentar una mayor proporción de estudiantes de familias con bajos ingresos,

estudiantes con discapacidades y estudiantes con necesidades de aprendizaje del inglés. Además, contaban con un porcentaje más elevado de alumnos y personal pertenecientes a grupos raciales y étnicos minoritarios, especialmente población negra y latina. Estos distritos también solían tener infraestructuras escolares más antiguas y en peor estado, incluyendo sistemas de ventilación menos eficientes y aulas más congestionadas.

Por otro lado, los distritos que retiraron las mascarillas en fechas más tempranas solían disponer de mejores condiciones estructurales, menor densidad de estudiantes por aula y perfiles sociodemográficos con menor representación de grupos tradicionalmente vulnerables. Este conjunto de diferencias refleja cómo factores socioeconómicos, demográficos y estructurales influyeron directamente en la toma de decisiones de cada distrito sobre el momento adecuado para levantar la medida de uso obligatorio de mascarillas (Cowger et al., 2022a).

Figura 3.3: Características de los Distritos Escolares



Fuente: elaboración propia del original en Cowger et al. (2022a).

### 3.2.1.2. Modelo de referencia

El modelo de referencia es el utilizado por Tori L. Cowger en Cowger et al. (2022a), propuesto por Callaway and Sant'Anna (2021b), que extiende el método clásico de diferencia en diferencias (Ashenfelter, 1978) para manejar situaciones en las que el tratamiento se adopta en distintos momentos. En cada periodo, los distritos tratados se comparan con aquellos que todavía no han recibido el tratamiento, permitiendo obtener comparaciones válidas en el tiempo.

El modelo se formaliza mediante la siguiente ecuación:

$$Y^{g,t} = \alpha_1^{g,t} + \alpha_2^{g,t} \cdot G_g + \alpha_3^{g,t} \cdot I\{T = t\} + \beta^{g,t} \cdot (G_g \times I\{T = t\}) + \epsilon^{g,t} \quad (3.1)$$

donde:  $Y^{g,t}$  es la tasa semanal de casos de COVID-19 por cada 100.000 estudiantes y/o personal;  $G_g$  es una variable indicadora que identifica a los distritos pertenecientes al grupo que recibe el tratamiento en la semana  $g$ ;  $I\{T = t\}$  es una variable indicadora para la semana de observación  $t$ ; el término  $\beta^{g,t}$  representa el efecto medio del tratamiento para el grupo  $g$  en el tiempo  $t$ , es decir, el parámetro de interés  $ATT(g,t)$  o *Average Treatment effect on the Treated*;  $\epsilon^{g,t}$  recoge el término de error (Cowger et al., 2022c).

Este planteamiento permite estimar de forma desagregada los efectos del tratamiento para cada grupo y cada instante temporal, adaptándose así a la variación escalonada en la implementación de la política de mascarillas. Además, el modelo incluye efectos fijos tanto a nivel de distrito como de semana, lo que permite controlar por factores invariables en el tiempo o comunes a todos los distritos en un periodo determinado.

A diferencia del diseño clásico de diferencia en diferencias (Ashenfelter, 1978), esta metodología no requiere suponer que el efecto del tratamiento es constante a lo largo del tiempo o entre grupos de tratamiento. Al estimar los efectos medios del tratamiento de forma separada para cada combinación de grupo y periodo, el modelo permite capturar posibles dinámicas heterogéneas en la respuesta al levantamiento de los mandatos de mascarilla.

### 3.2.1.3. Estimación del efecto

La estimación del efecto se basa en los parámetros  $\beta^{g,t}$  obtenidos a partir del modelo propuesto por Callaway and Sant'Anna (2021b). Estos parámetros recogen el efecto medio del tratamiento para el grupo  $g$  en la semana  $t$ , es decir, los  $ATT(g,t)$ .

Para obtener medidas agregadas del efecto, se utilizan los procedimientos utilizados en Cowger et al. (2022a): la agregación dinámica event-study y la agregación acumulada.

En primer lugar, el análisis dinámico permite representar la evolución temporal del efecto en función de las semanas transcurridas desde la retirada del mandato de mascarilla. Para cada semana relativa al tratamiento, se calcula un promedio ponderado de los efectos medios del tratamiento de los distintos grupos que han sido tratados en ese periodo. La fórmula empleada para esta agregación es:

$$\theta_c(t) = \sum_{g \in G} I\{t \geq g\} \cdot P(G = g | G \leq t) \cdot ATT(g, t) \quad (3.2)$$

donde  $G$  es el conjunto de grupos de tratamiento definidos por la semana de adopción,  $I\{t \geq g\}$  es una función indicadora que toma valor uno si el grupo  $g$  ya ha sido tratado

en el tiempo  $t$ , y  $P(G = g | G \leq t)$  representa la probabilidad condicional de pertenecer al grupo  $g$  entre los grupos que han recibido el tratamiento hasta el periodo  $t$  (Cowger et al., 2022c).

En segundo lugar, la estimación acumulada permite cuantificar el efecto total durante el conjunto del periodo de estudio, que en este caso abarca las 15 semanas posteriores al levantamiento del mandato estatal. La agregación acumulada se obtiene sumando los efectos dinámicos desde el inicio de la intervención hasta el periodo considerado, mediante la expresión:

$$\theta_{cum}(\tilde{t}) = \sum_{t=26}^{\tilde{t}} \theta_c(t) \quad (3.3)$$

donde  $t = 26$  corresponde a la primera semana posterior al levantamiento del mandato estatal, y  $\tilde{t}$  indica el periodo acumulado final (Cowger et al., 2022c).

Ambos procedimientos de agregación permiten obtener, respectivamente, una visión de la dinámica temporal del efecto del levantamiento de mascarillas y una medida global del exceso de casos atribuido a dicha medida.

Las ponderaciones utilizadas en las agregaciones reflejan el tamaño relativo de cada grupo de tratamiento, garantizando que las estimaciones representen fielmente la población de estudio. Además, los intervalos de confianza se obtienen mediante técnicas de bootstrap agrupadas a nivel de distrito, siguiendo la metodología propuesta por Callaway and Sant'Anna (2021b) e implementada en el plan de análisis estadístico del trabajo original.

Esta estrategia de estimación permite así cuantificar tanto la evolución temporal como el efecto acumulado del levantamiento escalonado de los mandatos de mascarilla en los distritos escolares analizados.

#### 3.2.1.4. Cumplimiento del supuesto de tendencias paralelas

El cumplimiento del supuesto de tendencias paralelas es necesario para el uso del modelo de diferencias en diferencias (Marx et al., 2024). Este supuesto establece que, en ausencia del tratamiento, la evolución de la incidencia de COVID-19 habría sido similar entre los distritos que levantaron el mandato de mascarillas y aquellos que aún no lo habían hecho. Bajo esta premisa, cualquier desviación observada tras la intervención puede atribuirse al efecto causal del tratamiento.

Para evaluar empíricamente este supuesto, en el apartado 4.1.1 se analizan los coeficientes  $ATT(g,t)$  estimados en los periodos previos al levantamiento de las mascarillas. Además, se aplicó el test estadístico de Cramér-von Mises (Anderson, 1962), que contrasta formalmente la igualdad de distribuciones acumuladas de los efectos pretratamiento frente a la hipótesis nula de tendencias paralelas.

El cumplimiento de este supuesto es fundamental para que las estimaciones obtenidas puedan interpretarse de forma causal, atribuyendo las diferencias observadas en la incidencia de casos al levantamiento escalonado de los mandatos de mascarilla.

### 3.2.1.5. Implementación en R

La implementación en R se ha llevado a cabo utilizando el paquete `did` (Callaway and Sant'Anna, 2021a). Este paquete permite aplicar de forma directa la metodología de diferencias en diferencias con tratamiento escalonado, estimando de manera separada los efectos medios del tratamiento para los tratados ( $ATT(g,t)$ ).

En primer lugar, se construyeron las variables necesarias para identificar el momento de tratamiento de cada distrito. La variable `first.treat` recoge la semana en la que cada distrito levantó el mandato de mascarilla, codificando los valores 26, 27 o 28 según la semana de adopción. Los distritos que no modificaron su política durante el periodo de estudio permanecen con el valor 0, indicando su condición de no tratados.

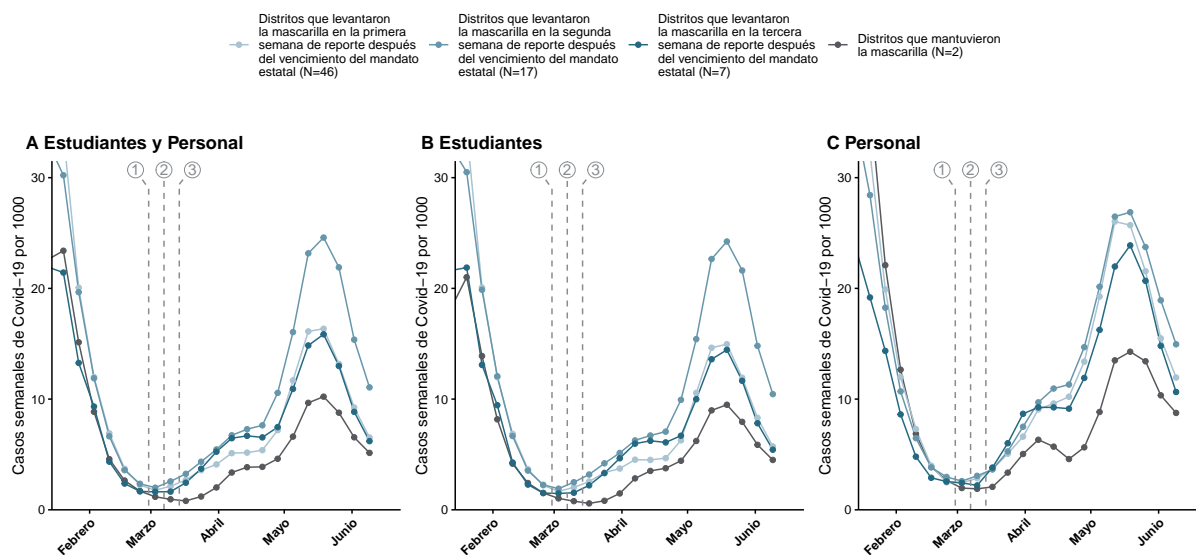
A continuación, se estimaron los efectos medios del tratamiento para los tratados mediante el estimador `group-time average treatment effect`, utilizando como grupo de control aquellos distritos que todavía no habían recibido el tratamiento en cada momento del tiempo (grupo `not-yet-treated`). Este enfoque garantiza que las comparaciones se realizan entre unidades comparables en cada periodo.

Posteriormente, se calcularon las agregaciones dinámicas (`event-study`) y las agregaciones por semana de calendario, obteniendo así tanto la evolución temporal del efecto como el efecto acumulado del levantamiento de las mascarillas. Las ponderaciones empleadas en estas agregaciones consideran el tamaño relativo de los distintos grupos de tratamiento en cada semana, reflejando así la estructura real de la población escolar analizada.

Finalmente, los intervalos de confianza de las estimaciones se calcularon mediante procedimientos de `bootstrap` agrupados a nivel de distrito, siguiendo las recomendaciones metodológicas establecidas por Callaway and Sant'Anna (2021b).

### 3.2.2. Estimaciones semanales y acumuladas del ATT

Figura 3.4: Incidencia de Covid-19 en los distritos escolares antes y después del levantamiento de las mascarillas



Fuente: elaboración propia del original en Cowger et al. (2022a).

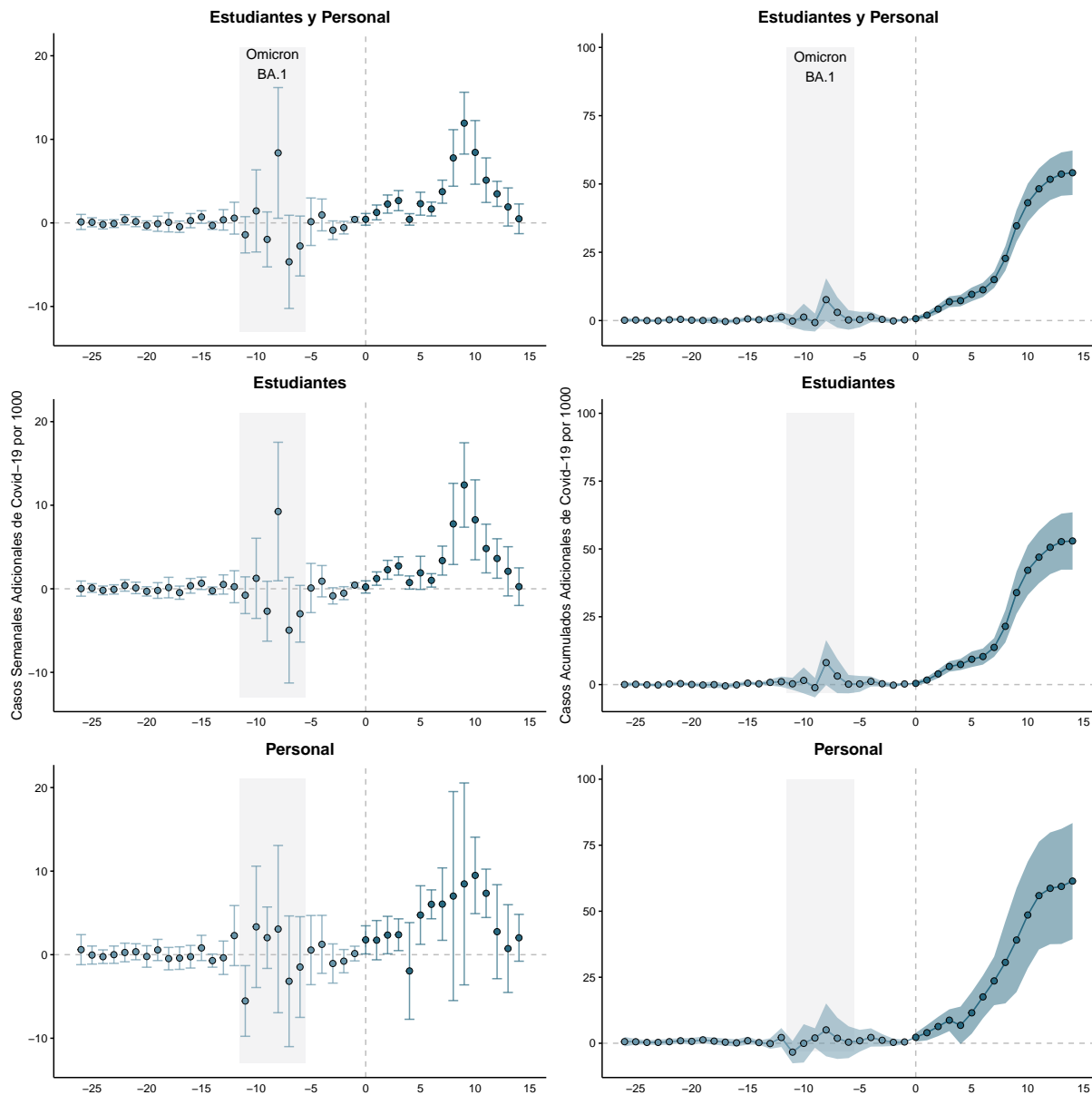
A partir de la metodología planteada en Cowger et al. (2022a) y descrita en los apartados anteriores, se obtienen las estimaciones dinámicas y acumuladas del efecto del levantamiento escalonado de los mandatos de mascarilla sobre la incidencia de casos de COVID-19 en los distritos escolares analizados.

En primer lugar, las estimaciones semanales muestran la evolución temporal del efecto tras la retirada de la obligatoriedad de mascarillas. Tal y como se observa en la primera columna de la Figura 3.5, el efecto comienza a ser evidente a partir de las primeras semanas tras la intervención, con un incremento progresivo en la incidencia de casos. Los valores estimados de los efectos medios semanales del tratamiento para los tratados presentan un patrón creciente durante las primeras semanas posteriores al tratamiento, acentuándose entre las semanas 8 y 10 en el caso de los estudiantes y el total. Por otra parte, se observa una mayor variabilidad en las estimaciones del personal, donde varias estimaciones incluyen el 0 en su intervalo de confianza.

En segundo lugar, las estimaciones acumuladas permiten cuantificar el impacto global de la retirada de mascarillas durante el conjunto del periodo de seguimiento. En este caso, se calcula el exceso acumulado de casos atribuible a la intervención a lo largo de las 15 semanas posteriores al levantamiento del mandato estatal. Los resultados acumulados, junto con sus intervalos de confianza, se presentan en la segunda columna Figura 3.5.



Figura 3.5: Estimaciones semanales y acumuladas del ATT por Diferencia en Diferencias



Fuente: elaboración propia del original en Cowger et al. (2022a).

Estas estimaciones reflejan el impacto total que tuvo la retirada de las mascarillas en la incidencia de casos de COVID-19 dentro de los centros educativos, proporcionando así una medida integrada de los efectos del tratamiento a lo largo del tiempo.

### 3.3. Extensión metodológica

#### 3.3.1. Generalized Synthetic Control

El método Generalized Synthetic Control (Xu and Liu, 2021) constituye una extensión del enfoque de control sintético clásico (Abadie and Gardeazabal, 2003) que permite estimar efectos causales en estudios observacionales de series temporales múltiples, incluso

cuando el supuesto de tendencias paralelas del modelo de diferencias en diferencias podría no cumplirse estrictamente.

A diferencia del modelo de diferencias en diferencias (Callaway and Sant'Anna, 2021b), que compara promedios de grupos, el generalized synthetic control construye para cada unidad tratada un contrafactual específico basado en una combinación ponderada de unidades no tratadas. Además, ajusta las trayectorias pasadas de cada unidad mediante factores latentes comunes y específicos de cada unidad y periodo.

### 3.3.1.1. Modelo de referencia

El modelo puede expresarse de la siguiente forma:

$$Y_{it} = \delta_{it}D_{it} + \lambda_i f_t + \epsilon_{it} \quad (3.4)$$

donde:  $Y_{it}$  es la variable de resultado observada para la unidad  $i$  en el periodo  $t$ ;  $D_{it}$  es la variable indicadora de tratamiento;  $\delta_{it}$  recoge el efecto del tratamiento para la unidad  $i$  en el periodo  $t$ ;  $\lambda_i$  son los factores específicos de cada unidad;  $f_t$  son los factores comunes no observados que evolucionan en el tiempo;  $\epsilon_{it}$  es el término de error (Xu and Liu, 2021).

### 3.3.1.2. Implementación en R

En la implementación realizada se ha empleado el paquete `gsynth` de R (Xu and Liu, 2021), aplicando el estimador de matriz completada (`mc`) que permite identificar los factores latentes mediante imputación iterativa.

La especificación incluye, además del indicador de tratamiento, variables de control que recogen los retardos de la variable de resultado hasta el segundo periodo anterior. De esta forma, se incorpora información dinámica que permite ajustar mejor las trayectorias previas de las unidades tratadas.

Se ha utilizado una descomposición de factores bidireccional (`force = "two-way"`) que permite capturar tanto efectos fijos por unidad como por periodo, junto a un rango de factores latentes  $r \in [0, 5]$ , seleccionado mediante validación cruzada automática (`CV = TRUE`). La estimación de los intervalos de confianza se ha llevado a cabo mediante técnicas de bootstrap no paramétrico agrupado, con 1000 repeticiones.

Esta implementación permite obtener estimaciones tanto de los efectos agregados como de los efectos dinámicos por semana, ofreciendo una alternativa robusta al modelo de diferencias en diferencias en contextos donde puedan existir diferencias estructurales o dinámicas complejas en la evolución temporal de las unidades.

## 3.3.2. Estimaciones semanales y acumuladas del ATT

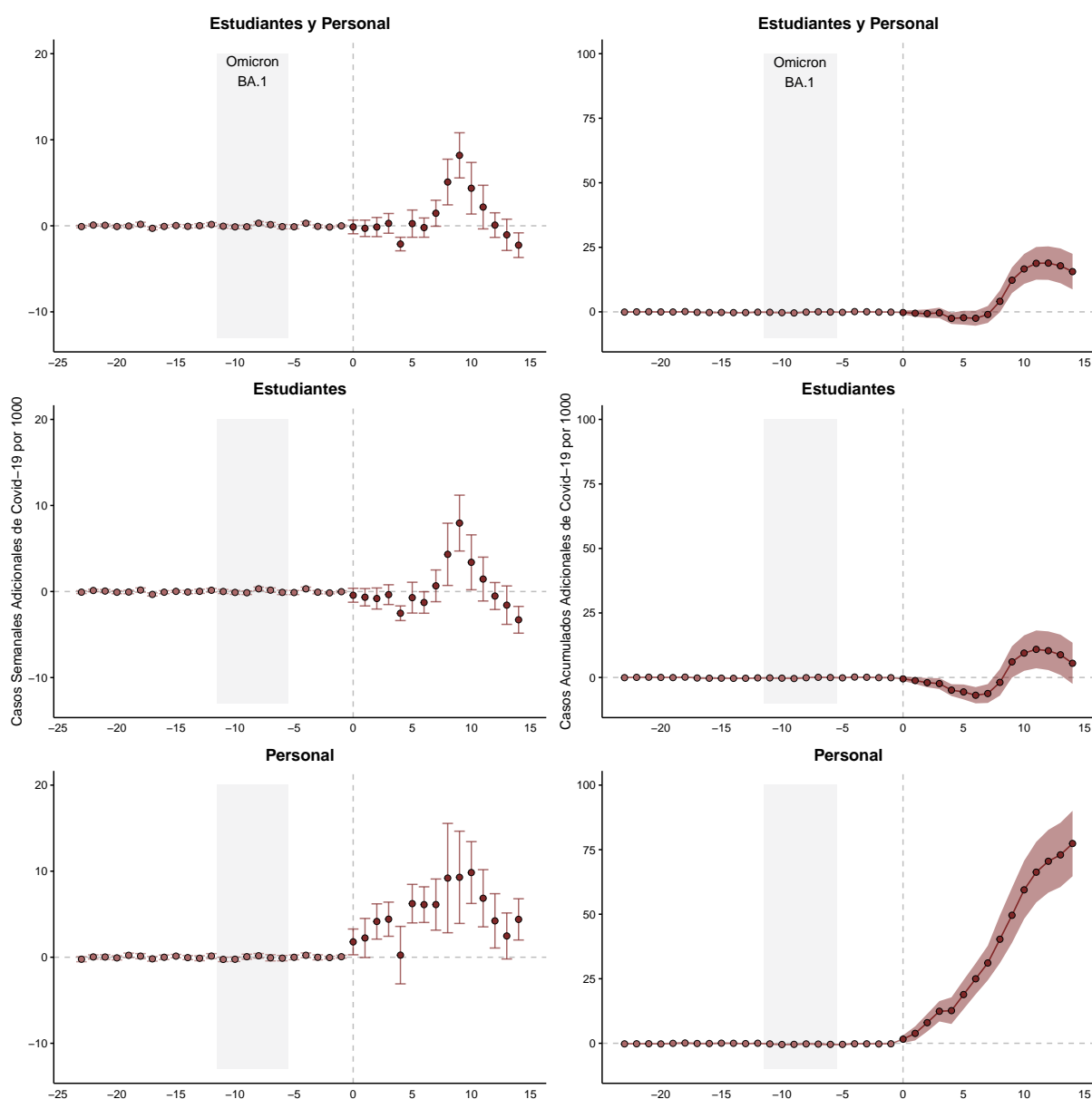
A partir de la estimación realizada mediante el método Generalized Synthetic Control, se obtuvieron las trayectorias temporales del efecto del levantamiento de los mandatos de mascarilla, tanto de forma semanal como acumulada, diferenciadas por población de estudio (estudiantes y personal).

En primer lugar, las estimaciones dinámicas muestran cómo la incidencia de casos de COVID-19 en estudiantes y en total presentan una estabilización alrededor del 0, siendo

sólo significativos entre las semanas 8 y 10. Por la parte del personal, se estima un mayor número de semanas significativas post-tratamiento.

En segundo lugar, las estimaciones acumuladas cuantifican el exceso total de casos atribuible al levantamiento de las mascarillas durante el conjunto del periodo de seguimiento.

Figura 3.6: Estimaciones semanales y acumuladas del ATT por Generalized Synthetic Control



Fuente: elaboración propia.

### 3.3.3. Comparación metodológica: ventajas, limitaciones y supuestos

La aplicación de ambos enfoques metodológicos permite contrastar las estimaciones obtenidas bajo diferentes supuestos y estructuras de modelización. A continuación, se presentan las principales ventajas, limitaciones y diferencias conceptuales entre el modelo

de diferencias en diferencias con adopción escalonada (Callaway and Sant'Anna, 2021b) y el método generalized synthetic control (Xu and Liu, 2021).

### **3.3.3.1. Supuesto de tendencias paralelas**

En primer lugar, el supuesto de tendencias paralelas es necesario para el uso del modelo de diferencia en diferencias escalonado (Marx et al., 2024): las trayectorias de los grupos tratados y no tratados habrían seguido una evolución similar en ausencia del tratamiento. Este supuesto es fundamental, aunque difícil de verificar completamente en contextos observacionales. En cambio, el método generalized synthetic control permite relajar esta condición al ajustar las trayectorias individuales mediante factores latentes, siendo capaz de acomodar posibles diferencias preexistentes en la evolución de las unidades antes de la intervención (Xu and Liu, 2021).

### **3.3.3.2. Flexibilidad temporal**

En segundo lugar, respecto a la flexibilidad temporal, el modelo generalized synthetic control permite capturar dinámicas heterogéneas en las trayectorias previas sin imponer restricciones específicas sobre la forma de la evolución pretratamiento (Xu and Liu, 2021). El modelo de diferencia en diferencias, por el contrario, asume que cualquier diferencia preexistente entre grupos es constante en el tiempo y queda absorbida por los efectos fijos (Callaway and Sant'Anna, 2021b).

### **3.3.3.3. Ventajas prácticas**

En tercer lugar, en cuanto a las ventajas prácticas, el método diferencia en diferencias escalonado presenta una interpretación directa de los efectos medios del tratamiento en los tratados, con estimaciones fácilmente desglosables por semanas relativas al tratamiento y ampliamente utilizadas en la literatura epidemiológica y de políticas públicas (Morbey et al., 2023). Por su parte, el generalized synthetic control proporciona estimaciones contrafactuales individualizadas para cada unidad, lo que permite una evaluación más detallada de los efectos a nivel de distrito, aunque a costa de una mayor complejidad computacional y una mayor sensibilidad a la especificación del modelo (Xu and Liu, 2021).

### **3.3.3.4. Limitaciones**

Por último, respecto a las limitaciones, el modelo generalized synthetic control requiere una cantidad suficiente de observaciones en el periodo pretratamiento para estimar correctamente los factores latentes (Xu and Liu, 2021), mientras que el modelo de diferencia en diferencias es más robusto en contextos con ventanas pretratamiento más reducidas (Callaway and Sant'Anna, 2021b). Por otro lado, el generalized synthetic control puede verse afectado por sobreajuste si el número de factores no es seleccionado correctamente (Xu and Liu, 2021), mientras que diferencia en diferencias depende más críticamente del cumplimiento del supuesto de tendencias paralelas (Marx et al., 2024).

# Capítulo 4

## Resultados

### 4.1. Resultados

En este apartado se presentan los resultados obtenidos tras la aplicación de las dos metodologías de inferencia causal analizadas: el modelo de Diferencias en Diferencias con tratamiento escalonado (Callaway and Sant'Anna, 2021b) utilizado en Cowger et al. (2022a) y el método Generalized Synthetic Control (Xu and Liu, 2021). Se muestran tanto las estimaciones acumuladas como los efectos globales obtenidos durante el periodo de 15 semanas posteriores al levantamiento de los mandatos de mascarilla.

Tabla 4.1: Casos Acumulados de Covid-19 durante el periodo de 15 semanas

Población	Todos los Distritos		Distritos que Levantaron la Mascarilla		Distritos que Mantuvieron la Mascarilla	
	Nº Casos	Nº Casos por 1000	Nº Casos	Nº Casos por 1000	Nº Casos	Nº Casos por 1000
Todos	40104	119.7	35337	134.1	4767	66.6
Estudiantes	31948	110.4	28273	123.7	3675	60.4
Personal	8156	178.3	7064	201.8	1092	101.6

Fuente: elaboración propia del original en Cowger et al. (2022a).

#### 4.1.1. Verificación del supuesto de tendencias paralelas

Antes de interpretar los efectos estimados mediante el modelo de diferencias en diferencias, es necesario verificar el cumplimiento del supuesto fundamental de tendencias paralelas (Marx et al., 2024). Este supuesto establece que, en ausencia del tratamiento, los distritos que levantaron el mandato de mascarillas y los que lo mantuvieron habrían seguido trayectorias similares en la evolución de los casos de COVID-19.

La validación de este supuesto se realizó a través de dos estrategias complementarias:

En primer lugar, se analizaron los coeficientes estimados en las semanas previas (tanto por diferencia en diferencias (Callaway and Sant'Anna, 2021b) como en generalized synthetic control (Xu and Liu, 2021)) al tratamiento mediante el análisis de eventos

(event-study). Los resultados muestran que las diferencias entre grupos en el periodo pre-tratamiento son, en su mayoría, próximas a cero y no estadísticamente significativas, lo que sugiere que no existían tendencias divergentes antes de la intervención.

En segundo lugar, se aplicó el test estadístico de Cramér–von Mises (Anderson, 1962), diseñado para evaluar formalmente la equivalencia en las distribuciones de los efectos pretratamiento. Este test no rechazó la hipótesis nula de igualdad de tendencias ( $p > 0.05$ ), lo que refuerza la validez del supuesto.

En conjunto, ambas evidencias permiten confirmar que el supuesto de tendencias paralelas se cumple razonablemente en el contexto de este estudio, lo que respalda la interpretación causal de los efectos estimados mediante el modelo de diferencias en diferencias con tratamiento escalonado.

#### 4.1.2. Resultados mediante Diferencia en Diferencias

Los resultados obtenidos con el modelo de diferencias en diferencias (Callaway and Sant'Anna, 2021b) muestran un incremento progresivo en la incidencia de casos tras la retirada de la obligatoriedad de mascarillas. El efecto comienza a manifestarse en las primeras semanas posteriores al tratamiento, presentando un patrón creciente especialmente a partir de la semana 8, donde se observa un mayor incremento en el caso de los estudiantes y el total de la población escolar. En el caso del personal, las estimaciones presentan mayor variabilidad, incluyendo algunos intervalos de confianza que abarcan el cero en determinadas semanas.

En términos acumulados, se estima que la retirada de las mascarillas generó 14.102 casos adicionales de COVID-19 en el conjunto de los distritos tratados durante las 15 semanas analizadas, lo que representa un exceso de 53,5 casos por cada 1000 personas (IC 95 %: 44,6 – 62,4). En el desglose por población, los estudiantes presentan 11.968 casos adicionales (52,4 por 1000; IC 95 %: 41,2 – 63,5), mientras que el personal presenta 2.125 casos adicionales (60,7 por 1000; IC 95 %: 37,8 – 83,6).

Tabla 4.2: Casos de Covid-19 estimados por Diferencia-en-Diferencias.

Población	Nº Casos Adicionales por 1000 (95 % CI)	Nº Casos Adicionales (95 % CI)	Porcentaje de los Casos en Distritos que Levantaron las Mascarillas (95 % CI)	Porcentaje de los Casos en Todos los Distritos (95 % CI)
Todos	53.5 (44.4-62.7)	14102 (11691-16512)	39.9 (33.1-46.7)	35.2 (29.2-41.2)
Estudiantes	52.4 (40.6-64.1)	11968 (9285-14652)	42.3 (32.8-51.8)	37.5 (29.1-45.9)
Personal	60.7 (37-84.4)	2125 (1294-2955)	30.1 (18.3-41.8)	26.1 (15.9-36.2)

Fuente: elaboración propia del original en Cowger et al. (2022a).

### 4.1.3. Resultados mediante Generalized Synthetic Control

El modelo generalized synthetic control (Xu and Liu, 2021) arroja estimaciones de menor magnitud global, pero confirma la existencia de un efecto relevante, particularmente en el personal escolar.

Las estimaciones dinámicas muestran que, en estudiantes y en el conjunto total, los efectos semanales se mantienen mayoritariamente próximos a cero durante gran parte del periodo posterior al tratamiento, alcanzando significación estadística únicamente entre las semanas 8 y 10. Por el contrario, en el caso del personal, el número de semanas con efectos significativos es superior, indicando una mayor sensibilidad de este colectivo al levantamiento de las mascarillas.

En cuanto a los efectos acumulados, se estiman 4.107 casos adicionales en el conjunto de distritos tratados, lo que supone un exceso de 15,6 casos por 1000 personas (IC 95 %: 8,6 – 22,6). Por grupos, los estudiantes presentan un exceso de 1.316 casos (5,8 por 1000; IC 95 %: -2,2 – 13,7), mientras que el personal acumula 2.709 casos adicionales (77,4 por 1000; IC 95 %: 64,8 – 90,0), siendo este último el grupo donde el efecto resulta más evidente.

Tabla 4.3: Casos de Covid-19 estimados por Generalized Synthetic Control.

Población	Nº Casos Adicionales por 1000 (95 % CI)	Nº Casos Adicionales (95 % CI)	Porcentaje de los Casos en Distritos que Levantaron las Mascarillas (95 % CI)	Porcentaje de los Casos en Todos los Distritos (95 % CI)
Todos	15.6 (8.6-22.5)	4107 (2275-5938)	11.6 (6.4-16.8)	10.2 (5.7-14.8)
Estudiantes	5.5 (-2.5-13.5)	1255 (-579-3090)	4.4 (-2-10.9)	3.9 (-1.8-9.7)
Personal	77.4 (64.6-90.1)	2709 (2263-3155)	38.3 (32-44.7)	33.2 (27.7-38.7)

Fuente: elaboración propia.

## 4.2. Discusión

Las estimaciones obtenidas en este trabajo consiguen replicar parcialmente los efectos reportados en Cowger et al. (2022a), lo que sugiere la posible influencia de diferencias en el tratamiento de los datos, ajustes metodológicos específicos o criterios de inclusión y exclusión de distritos.

Por otro lado, ambos enfoques metodológicos muestran efectos consistentes en términos de dirección del impacto, indicando un aumento de los casos de COVID-19 tras la retirada de las mascarillas. No obstante, la magnitud de los efectos difiere sustancialmente entre ambos métodos. Mientras que el modelo de diferencia en diferencias (Callaway and Sant'Anna, 2021b) estima un impacto general elevado, tanto en estudiantes como en personal, el método generalized synthetic control (Xu and Liu, 2021) detecta un efecto más

atenuado, centrado fundamentalmente en el personal escolar, mientras que en estudiantes los efectos acumulados resultan más modestos y no siempre estadísticamente significativos.

Estas diferencias metodológicas reflejan la distinta sensibilidad de cada modelo ante las posibles desviaciones de los supuestos de identificación, tal y como se ha analizado previamente en la comparación metodológica. Pese a ello, el patrón general de resultados es coherente: la retirada de los mandatos de mascarilla se asocia con un aumento en la incidencia de casos, especialmente entre el personal escolar, población potencialmente más vulnerable dentro del ámbito educativo.



# Capítulo 5

## Conclusiones

### 5.1. Conclusiones

Este trabajo ha tenido como objetivo principal reproducir la metodología del estudio Cowger et al. (2022a) a partir de la documentación complementaria facilitada por los autores (Cowger et al., 2022c) (Cowger et al., 2022b). De este modo, se implementa un modelo propio de diferencia en diferencias con tratamiento escalonado (Callaway and Sant'Anna, 2021b), y se introduce además un segundo enfoque metodológico alternativo, el *generalized synthetic control* (Xu and Liu, 2021), con el fin de comparar los resultados y evaluar la robustez de las conclusiones bajo diferentes supuestos estadísticos (Baker et al., 2022).

Los resultados obtenidos con ambas metodologías indican que la retirada de las mascarillas se asoció con un aumento de los casos de COVID-19 en los distritos escolares analizados. No obstante, las estimaciones obtenidas en este trabajo replican parcialmente los efectos reportados en Cowger et al. (2022a), lo que sugiere la posible influencia de diferencias en el tratamiento de los datos, ajustes metodológicos específicos o criterios de inclusión y exclusión de distritos. Pese a ello, el patrón general es consistente: se observa un exceso de casos en los distritos que levantaron la obligatoriedad del uso de mascarillas frente a aquellos que la mantuvieron.

El modelo de diferencia en diferencias (Callaway and Sant'Anna, 2021b) estima un exceso de más de 14.000 casos durante el periodo de seguimiento, afectando tanto a estudiantes como a personal. Por su parte, el método *generalized synthetic control* (Xu and Liu, 2021) confirma un efecto significativo especialmente en el personal escolar, aunque con estimaciones globales más moderadas, y efectos no significativos para los estudiantes.

La aplicación de ambos enfoques ha permitido contrastar los resultados bajo diferentes supuestos metodológicos, fortaleciendo así la validez del análisis. Mientras que el modelo diferencia en diferencias (Callaway and Sant'Anna, 2021b) se apoya en el cumplimiento del supuesto de tendencias paralelas (Marx et al., 2024) —verificado mediante análisis gráfico y pruebas estadísticas—, el modelo *generalized synthetic control* (Xu and Liu, 2021) ofrece una alternativa más flexible para ajustar dinámicas pretratamiento heterogéneas entre distritos.

En conjunto, aunque los resultados obtenidos no replican exactamente los de Cowger et al. (2022a), este trabajo constituye una reproducción razonada, técnicamente sólida y metodológicamente extensible, cuyo objetivo final es contribuir a la comprensión del

fenómeno desde una perspectiva rigurosa y transparente.

## 5.2. Trabajo futuro

El análisis desarrollado en este trabajo ha permitido reproducir y extender, mediante diferentes enfoques metodológicos, los resultados del estudio Cowger et al. (2022a) sobre el impacto del levantamiento de las mascarillas en centros escolares. No obstante, existen múltiples vías para seguir profundizando y mejorando este tipo de análisis en trabajos futuros.

En primer lugar, sería conveniente incorporar variables adicionales de control que puedan capturar diferencias estructurales entre distritos, como el tamaño medio de las aulas, las condiciones de ventilación, o la densidad de estudiantes por metro cuadrado (Weaver et al., 2022). Estas variables podrían incluirse tanto en el modelo de diferencias en diferencias (Callaway and Sant'Anna, 2021b) como en el modelo *generalized synthetic control* (Xu and Liu, 2021), mejorando la especificación y reduciendo posibles sesgos por omisión.

En segundo lugar, podría explorarse el uso de métodos que incorporen emparejamiento previo al análisis causal, como el *matching* por puntuación de propensión (Ho et al., 2011) o el emparejamiento exacto (King and Nielsen, 2019), con el fin de reforzar la comparabilidad entre distritos tratados y no tratados antes de aplicar los modelos de estimación.

Otra línea de trabajo sería el análisis del impacto diferenciado según características sociodemográficas, como el nivel de renta del distrito, la proporción de estudiantes con necesidades especiales o la prevalencia de población vulnerable (Karmakar et al., 2021). Esto permitiría evaluar si el efecto del levantamiento de las mascarillas fue homogéneo o si afectó de forma desigual a distintos grupos.

Además, sería interesante evaluar la robustez de los resultados utilizando otras variables de resultado, como el absentismo escolar, el número de días perdidos por enfermedad o las hospitalizaciones relacionadas con COVID-19. Estas métricas podrían aportar una perspectiva más amplia del impacto de la retirada de mascarillas, más allá de los casos confirmados.

Por último, se podrían explorar métodos adicionales de inferencia causal como el *augmented synthetic control* (Ben-Michael et al., 2021), que podría mejorar la estimación en contextos con alta heterogeneidad entre unidades y ventanas temporales limitadas.

En resumen, estas posibles líneas de trabajo permitirían no sólo mejorar la precisión de las estimaciones, sino también extender la aplicabilidad del análisis a otros contextos educativos, epidemiológicos o de formulación de políticas públicas.

# Apéndice A

## Anexos

### A.1. Código

Listado A.1: Función R para leer los reportes PDF de COVID-19

```
obt_data_distritos <- function(archivos){
  fechas <- str_extract(string = archivos,
                        pattern = "\\d{4}-\\d{1,2}-\\d{1,2}")
  df <- data.frame(Code = NA, Name = NA, Students = NA,
                  Staff = NA, PT = NA, PPT = NA,
                  PPR = NA, Date = NA)
  for(i in 1:length(archivos)){
    df2 <- pdf_text(paste0("data/reportes_covid/",archivos[i]))
    p <- 2
    fin <- FALSE
    while(fin == FALSE){
      if(p == 2){
        aux <- df2[[p]] %>%
          str_split("\\n") %>%
          as_tibble(.name_repair=make.names) %>%
          slice(7:nrow(.)) %>%
          separate(X, into = c("Code", "Name", "Students", "Staff",
                              "PT", "PPT", "PPR"), sep = "\\s{2,}") %>%
          filter(Code != "") %>%
          mutate(Date = fechas[i])
      } else{
        aux <- df2[[p]] %>%
          str_split("\\n") %>%
          as_tibble(.name_repair=make.names) %>%
          separate(X, into = c("Code", "Name", "Students", "Staff",
                              "PT", "PPT", "PPR"), sep = "\\s{2,}") %>%
          filter(Code != "") %>%
          mutate(Date = fechas[i])
      }
      df <- rbind(df, aux)
      p <- p + 1
      fin <- grepl(pattern = "Education Collaboratives", df2[[p]])
    }
  }
  return(df)
}
```

## Listado A.2: Función R para corrección de semanas con reportes quincenales

```

week2correction <- function(df_raw){
  df_selected <- df_raw %>%
    select(Code, Name, Week_lifted, Date, Students, Staff, PT, PPT)

  df_list <- list()

  for(code in unique(df_selected$Code)){
    aux <- df_selected %>%
      filter(Code == code) %>%
      arrange(Date)

    if(nrow(aux) > 0){
      temp_rows <- tibble(
        Code = character(),
        Name = character(),
        Week_lifted = character(),
        Date = as.Date(character()),
        Students = numeric(),
        Staff = numeric(),
        PT = numeric(),
        PPT = numeric()
      )

      i <- 1
      while(i <= nrow(aux)){
        if(i < nrow(aux) & as.numeric(aux$Date[i+1] - aux$Date[i]) == 14){
          missing_week_date <- aux$Date[i] + 7
          students_split <- aux$Students[i+1] / 2
          staff_split <- aux$Staff[i+1] / 2

          temp_rows <- bind_rows(temp_rows,
            tibble(Code = aux$Code[i],
              Name = aux$Name[i],
              Week_lifted = aux$Week_lifted[i],
              Date = missing_week_date,
              Students = students_split,
              Staff = staff_split,
              PT = 0,
              PPT = 0 ))

          aux$Students[i+1] <- students_split
          aux$Staff[i+1] <- staff_split

          temp_rows <- bind_rows(temp_rows,
            aux[i,] %>%
              select(Code, Name, Week_lifted, Date, Students, Staff,
                PT, PPT))
        } else{
          temp_rows <- bind_rows(temp_rows,
            aux[i,] %>%
              select(Code, Name, Week_lifted, Date, Students, Staff,
                PT, PPT))
        }
        i <- i + 1
      }

      df_list[[code]] <- temp_rows %>% arrange(Date)
    }
  }
}

```

```

df_week2corrected <- bind_rows(df_list)
df_week2corrected <- df_week2corrected %>% mutate(All = Students + Staff)
}

return(df_week2corrected)
}

```

Listado A.3: Función R para corrección de ceros no informativos en registros semanales

```

cases0correction <- function(df_week2corrected){
  df_cases0correction <- df_week2corrected %>%
    arrange(Code, Date) %>%
    group_by(Code) %>%
    mutate(ma = slide_dbl(All,
                          ~ mean(., na.rm = TRUE),
                          .before = 4, .after = -1, .complete = TRUE)) %>%
    ungroup()

  df_cases0correction <- df_cases0correction %>%
    mutate(is_zero_cases = All == 0,
           is_non_reporting_zero = is_zero_cases & (PPT >= 1 | (!is.na(ma) & ma > 10)),
           na_prev = FALSE)

  df_cases0correction_list <- list()

  for(code in unique(df_cases0correction$Code)){
    aux <- df_cases0correction %>%
      filter(Code == code) %>%
      select(Code, Name, Week_lifted, Date, Students, Staff, PT, PPT,
             All, is_non_reporting_zero, is_zero_cases, ma, na_prev) %>%
      arrange(Date)

    i <- 1
    while(i <= nrow(aux)){
      if(aux$is_non_reporting_zero[i]){
        aux$All[i] <- NA
        aux$Students[i] <- NA
        aux$Staff[i] <- NA
        if(i < nrow(aux)){
          aux$na_prev[i+1] <- TRUE
        }
      }
      i <- i + 1
    }
    df_cases0correction_list[[code]] <- aux
  }

  df_cases0correction <- bind_rows(df_cases0correction_list)
  df_cases0correction <- df_cases0correction %>%
    select(Code, Name, Week_lifted, Date, Students, Staff, All,
           is_non_reporting_zero, is_zero_cases, ma, na_prev)

  return(df_cases0correction)
}

```

Listado A.4: Función R para calcular tasas por 1.000 personas y marcar semanas sin mascarilla

```
weightcorrection <- function(df_cleaned){
  enrollmentbygrade <- read_excel("data/enrollmentbygrade.xlsx", skip = 1)
  staffracegender <- read_excel("data/staffracegender.xlsx", skip = 1)

  enrollment_data <- enrollmentbygrade %>%
    rename(Code = 'District Code', N_Students = Total) %>%
    select(Code, N_Students) %>%
    mutate(N_Students = as.numeric(str_remove_all(N_Students, ","))) %>%
    filter(!is.na(Code))

  staffing_data <- staffracegender %>%
    rename(Code = 'District/School Code', N_Staff = 'FTE Count') %>%
    select(Code, N_Staff) %>%
    mutate(N_Staff = as.numeric(str_remove_all(N_Staff, ","))) %>%
    filter(!is.na(Code))

  df_weighted <- df_cleaned %>%
    filter(!is.na(All)) %>%
    left_join(enrollment_data, by = "Code") %>%
    left_join(staffing_data, by = "Code") %>%
    mutate(
      N_Students = ifelse(is.na(N_Students), 2 * N_Students, N_Students),
      N_Staff = ifelse(is.na(N_Staff), 2 * N_Staff, N_Staff),
      N_All = N_Students + N_Staff,
      All_1000 = (Students + Staff) / N_All * 1000,
      Students_1000 = Students / N_Students * 1000,
      Staff_1000 = Staff / N_Staff * 1000,
      Week = floor(as.numeric(Date - as.Date("2021-09-9"))) / 7,
      mask_lifted = case_when(
        Week_lifted == 1 & Week >= 26 ~ TRUE,
        Week_lifted == 2 & Week >= 27 ~ TRUE,
        Week_lifted == 3 & Week >= 28 ~ TRUE,
        .default = FALSE
      )
    ) %>%
    select(Code, Name, N_All, N_Students, N_Staff, Week_lifted, Date,
           Week, mask_lifted, All, Students, Staff,
           All_1000, Students_1000, Staff_1000)

  return(df_weighted)
}
```

# Bibliografía

- Abadie, A. and Gardeazabal, J. (2003). The economic costs of conflict: A case study of the basque country. *American Economic Review*, 93.
- Anderson, T. W. (1962). On the distribution of the two-sample cramer-von mises criterion. *The Annals of Mathematical Statistics*, 33.
- Andrejko, K. L., Pry, J. M., Myers, J. F., Fukui, N., DeGuzman, J. L., Openshaw, J., Watt, J. P., Lewnard, J. A., Jain, S., Abdulrahim, Y., Barbaduomo, C. M., Bermejo, M. I., Cheunkarndee, J., Cornejo, A. F., Corredor, S., Dabbagh, N., Dong, Z. N., Dyke, A., Fang, A. T., Felipe, D., Frost, P. M., Ho, T., Javadi, M. H., Kaur, A., Lam, A., Li, S. S., Miller, M., Ni, J., Park, H., Poindexter, D. J., Samani, H., Saretha, S., Spencer, M., Spinoso, M. M., Tran, V. H., Walas, N., Wan, C., and Xavier, E. (2022). Effectiveness of face mask or respirator use in indoor public settings for prevention of sars-cov-2 infection — california, february–december 2021. *MMWR. Morbidity and Mortality Weekly Report*, 71.
- Ashenfelter, O. (1978). Estimating the effect of training programs on earnings. *The Review of Economics and Statistics*, 60.
- Baker, A. C., Larcker, D. F., and Wang, C. C. (2022). How much should we trust staggered difference-in-differences estimates? *Journal of Financial Economics*, 144.
- Ben-Michael, E., Feller, A., and and, J. R. (2021). The augmented synthetic control method. *Journal of the American Statistical Association*, 116(536):1789–1803.
- Boston Herald (2022). Massachusetts to Lift School Mask Mandate on February 28. <https://www.bostonherald.com/2022/02/09/massachusetts-lifting-school-mask-mandate-february-28/>. Published February 9, 2022.
- Boston Public Schools (2021). Boston Public Schools District Calendar: School Year 2021–2022. [https://www.bostonpublicschools.org/cms/lib/MA01906464/Centricity/Domain/4/BPS%20Cal%20SY22\\_FINAL.pdf](https://www.bostonpublicschools.org/cms/lib/MA01906464/Centricity/Domain/4/BPS%20Cal%20SY22_FINAL.pdf). Published 2021; PDF calendar subject to change due to extenuating circumstances or inclement weather.
- Branch-Elliman, W., Perkins, E., and Doron, S. (2022). Why Lifting Mask Mandates Is So Divisive. <https://commonwealthbeacon.org/opinion/why-lifting-mask-mandates-is-so-divisive/>. Published February 26, 2022; republished/updated September 22, 2023.
- Brookline Town News (2022). Town of Brookline to Reinstate Mask Mandate in Indoor Town-Owned Spaces Where Public Gathers. <https://brooklinetownnews.com/2022/05/20/town-of-brookline-to-reinstate-mask-mandate-in-indoor-town-owned-spaces-where-public-gathers/>. Published May 20, 2022; mandate effective May 23, 2022.

- Callaway, B. and Sant'Anna, P. H. (2021a). *did: Difference in Differences*. R package version 2.1.2.
- Callaway, B. and Sant'Anna, P. H. (2021b). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225.
- Cowger, T. L., Murray, E. J., Clarke, J., Bassett, M. T., Ojikutu, B. O., Sánchez, S. M., Linos, N., and Hall, K. T. (2022a). Lifting universal masking in schools — covid-19 incidence among students and staff. *New England Journal of Medicine*, 387.
- Cowger, T. L., Murray, E. J., Clarke, J., Bassett, M. T., Ojikutu, B. O., Sánchez, S. M., Linos, N., and Hall, K. T. (2022b). Statistical analysis plan for “Lifting universal masking in schools — covid-19 incidence among students and staff”.
- Cowger, T. L., Murray, E. J., Clarke, J., Bassett, M. T., Ojikutu, B. O., Sánchez, S. M., Linos, N., and Hall, K. T. (2022c). Supplement to “Lifting universal masking in schools — covid-19 incidence among students and staff”.
- Decker, S. (2021). Which States Banned Mask Mandates in Schools, and Which Required Masks? <https://www.edweek.org/policy-politics/which-states-ban-mask-mandates-in-schools-and-which-require-masks/2021/08>. Published August 20, 2021; last updated July 8, 2022.
- Gettings, J., Czarnik, M., Morris, E., Haller, E., Thompson-Paul, A. M., Rasberry, C., Lanzieri, T. M., Smith-Grant, J., Aholou, T. M., Thomas, E., Drenzek, C., and MacKellar, D. (2021). Mask use and ventilation improvements to reduce covid-19 incidence in elementary schools — georgia, november 16–december 11, 2020. *MMWR. Morbidity and Mortality Weekly Report*, 70.
- Ho, D. E., Imai, K., King, G., and Stuart, E. A. (2011). MatchIt: Nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software*, 42(8):1–28.
- Karmakar, M., Lantz, P. M., and Tipirneni, R. (2021). Association of Social and Demographic Factors With COVID-19 Incidence and Death Rates in the US. *JAMA Network Open*, 4(1):e2036462.
- King, G. and Nielsen, R. (2019). Why propensity scores should not be used for matching. *Political Analysis*, 27(4):435–454.
- Marx, P., Tamer, E., and Tang, X. (2024). Parallel trends and dynamic choices. *Journal of Political Economy Microeconomics*, 2.
- Masks4All (2020). Cloth masks can help stop the spread of covid-19, save lives and restore jobs. <https://web.archive.org/web/20200925214642/https://masks4all.co/>. Archived from the original on 25 September 2020. Retrieved 27 September 2020.
- Massachusetts Department of Elementary and Secondary Education (2023). Positive COVID-19 Cases in Massachusetts School Districts. <https://archives.lib.state.ma.us/collections/ff084038-2e48-47ef-8bd0-6b6e3281fe3c>. Data reported by the Massachusetts Department of Elementary and Secondary Education; accessed from Massachusetts State Archives digital collection.



- Massachusetts Department of Elementary and Secondary Education (2024). Statewide Accountability Report — School and District Profiles. <https://profiles.doe.mass.edu/statereport/#Assessment%20and%20Accountability>. Data last updated October 29, 2024.
- Morbey, R., Todkill, D., Deangelis, D., Charlett, A., and Elliot, A. (2023). Did it?: a differences-in-differences investigation tool to quantify the impact of local incidents on public health using real-time syndromic surveillance health data. *Epidemiology and Infection*, 151.
- Ooms, J. (2024). *pdftools: Text Extraction, Rendering and Converting of PDF Documents*. R package version 3.4.1.
- ProximityOne (2013). Boston–cambridge–newton, ma–nh msa situation & outlook report. <https://proximityone.com/metros/2013/cbsa14460.htm>. Published 2013; includes annual demographic and economic estimates and projections.
- Talic, S., Shah, S., Wild, H., Gasevic, D., Maharaj, A., Ademi, Z., Li, X., Xu, W., Mesa-Eguiagaray, I., Rostron, J., Theodoratou, E., Zhang, X., Motee, A., Liew, D., and Ilic, D. (2021). Effectiveness of public health measures in reducing the incidence of covid-19, sars-cov-2 transmission, and covid-19 mortality: systematic review and meta-analysis. *BMJ*, 375.
- Weaver, A. K., Head, J. R., Gould, C. F., Carlton, E. J., and Remais, J. V. (2022). Environmental factors influencing covid-19 incidence and severity. *Annual Review of Public Health*, 43.
- Xu, Y. and Liu, L. (2021). *gsynth: Generalized Synthetic Control Method*. R package version 1.2.1.