



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

ELECCIONES EN PANDEMIA: EFECTOS EN LA COMPOSICIÓN DEMOGRÁFICA DE LOS VOTANTES Y EN LAS PREFERENCIAS DE VOTO

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERA CIVIL INDUSTRIAL

CAMILA CONDE RODRÍGUEZ

PROFESOR GUÍA:
MARCELO OLIVARES ACUÑA

MIEMBROS DE LA COMISIÓN:
ALEJANDRA PUENTE CHANDIA
ANDRÉS MUSALEM

SANTIAGO DE CHILE

2024

RESUMEN DE LA MEMORIA PARA OPTAR
AL TÍTULO DE INGENIERÍA CIVIL INDUSTRIAL
POR: CAMILA CONDE RODRÍGUEZ
FECHA: 2024
PROF. GUÍA: MARCELO OLIVARES ACUÑA

ELECCIONES EN PANDEMIA: EFECTOS EN LA COMPOSICIÓN DEMOGRÁFICA DE LOS VOTANTES Y EN LAS PREFERENCIAS DE VOTO

El presente trabajo de investigación tiene como objetivo analizar los efectos de factores observables asociados a COVID-19 en las elecciones políticas nacionales que se realizaron durante el periodo de pandemia en Chile: el plebiscito nacional de 2020 y las elecciones presidenciales de 2021. Se estudian particularmente los efectos de la incidencia del virus, y el tiempo en las fases más estrictas del Plan Paso a Paso, en la composición demográfica de los votantes y en las preferencias de voto. Utilizando los resultados de las elecciones publicados por el Servicio Electoral a nivel de mesa de votación, junto a los métodos de regresiones de diferencia-en-diferencias (*DiD*) y modelos *logit* multinomiales, se estudian las siguientes hipótesis planteadas: (1a) periodos prolongados de confinamiento fueron asociados a una mayor proporción de votantes jóvenes, (1b) altos niveles de incidencia fueron asociados a una menor proporción de votantes mayores en edad, y (2) la incidencia de COVID-19 y el tiempo en confinamiento incidieron en cambios de preferencia de voto por parte de votantes ubicados en zonas con altos niveles de incidencia y períodos más prolongados de confinamiento.

Lo anterior es motivado por efectos del COVID-19 observados en otros países como en Francia, donde Giommoni and Loumeau [2020] mostraron que el confinamiento incentivó a electores en zonas con mayores restricciones a movilizarse y por ende participar en las elecciones municipales. Adicionalmente, Scheller [2021] encontró que electores mayores en edad y en zonas con mayores tasas de contagio en el estado de Florida fueron menos probables de votar en las elecciones presidenciales de Estados Unidos. A partir de los resultados de los modelos estimados para las elecciones chilenas, los estimadores *DiD* indican que las mesas de votación ubicadas en zonas con alta incidencia de COVID-19 se asociaron con una menor proporción de votantes entre 60-69 años durante el plebiscito nacional de 2020 en comparación a mesas ubicadas en comunas con baja incidencia. Además, zonas bajo un mayor tiempo en confinamiento estricto tuvieron efectos positivos en la proporción de votantes entre 18-29 y 40-49 años en el plebiscito nacional. Estos resultados implican que no hay suficiente evidencia para respaldar lo planteado en las hipótesis *H1a* y *H1b* para ambos el plebiscito nacional y la segunda vuelta presidencial.

Finalmente, al modelar mediante *logit* multinomiales las preferencias de voto en la primera vuelta presidencial de 2021 a nivel de mesa de votación, los parámetros estimados indican que las mesas con votantes iguales o similares a la primera vuelta presidencial anterior, y que previamente apoyaron mayormente a Sebastian Piñera en 2017, fueron más probables de votar por José Kast por sobre Gabriel Boric en comparación a las mesas con nuevos votantes. Se obtienen los mismos resultados incorporando los factores de incidencia y confinamiento, demostrando que estos factores no incidieron en cambios de preferencia durante la primera vuelta presidencial en pandemia y por lo tanto refutando la hipótesis *H2*.

Tabla de Contenido

1. INTRODUCCIÓN	1
1.1. El Sistema Electoral de Chile	1
1.2. Pandemia en Chile	2
1.3. Literatura	8
1.4. Objetivos	9
1.5. Alcance	9
2. MARCO TEÓRICO	10
2.1. Métodos	10
2.1.1. Regresión lineal multiple	10
2.1.2. Regresión logística multiple o Logit binomial	10
2.1.3. Diferencia-en-Diferencias (DiD)	11
2.1.4. Logit multinomial	12
3. DESARROLLO METODOLÓGICO	13
3.1. Metodología <i>CRISP-DM</i>	13
3.2. Entendimiento del problema	14
3.3. Entendimiento de los datos	15
3.3.1. Descripción de datos	15
3.3.2. Análisis exploratorio de datos	16
3.4. Preparación de datos	28
3.5. Modelamiento	29
3.5.1. <i>DiD</i>	29
3.5.2. <i>Logit</i> multinomial	32
3.6. Evaluación	37
4. CONCLUSIONES	39
Bibliografía	41
Anexo A.1	42
Anexo A.2	43
Anexo A.3	43
Anexo A.4	43
Anexo A.5	44
Anexo A.6	45
Anexo A.7	46
Anexo A.8	46
Anexo A.9	47

Anexo A.10	47
----------------------	----

Índice de Tablas

3.1.	Definición de variables	15
3.2.	Estadísticas descriptivas: características mesas y comunas correspondientes al plebiscito nacional de 2020	17
3.3.	Estadísticas descriptivas: resultados primera vuelta presidencial de 2021 y características a nivel de mesa y comuna	18
3.4.	Estadísticas descriptivas: características mesas y comunas correspondientes a la segunda vuelta presidencial de 2021	19
3.5.	Resumen de prueba Kruskal-Wallis para testear por diferencias significativas de cambios en participación entre grupos etarios	25
3.6.	Resultados de pruebas t de Welch para comparar medias de cambios en participación en el plebiscito nacional entre comunas en base a <i>Fases12</i>	25
3.7.	Pruebas <i>Wilcoxon Rank Sum</i> para comparar cambios en participación en el plebiscito nacional entre comunas en base a <i>Fases12</i>	25
3.8.	Resultados de pruebas t de Welch para comparar medias de cambios en participación en el plebiscito nacional entre comunas en base a <i>Incidencia</i>	25
3.9.	Pruebas <i>Wilcoxon Rank Sum</i> para comparar cambios en participación en el plebiscito nacional entre comunas en base a <i>Incidencia</i>	26
3.10.	Conteo de mesas ganadas por candidato utilizando resultados de elecciones presidenciales publicados por el Servel (2017 vs 2021)	27
3.11.	Resultados de estimación <i>DiD</i> para determinar significancia de <i>Fases12</i> en la proporción de votantes por grupo etario en el plebiscito nacional	30
3.12.	Resultados de estimación <i>DiD</i> para determinar significancia de <i>Incidencia</i> en la proporción de votantes por grupo etario en el plebiscito nacional	31
3.13.	Resumen de logit multinomial de la primera vuelta presidencial de 2021 (ML1)	33
3.14.	Resumen de logit multinomial de la primera vuelta presidencial de 2021 (ML2)	34
3.16.	Resumen de logit multinomial de la primera vuelta presidencial de 2021 (ML4)	34
3.15.	Resumen de logit multinomial de la primera vuelta presidencial de 2021 (ML3)	35
3.17.	Resumen de logit multinomial de la primera vuelta presidencial de 2021 (ML5)	36
3.18.	Valores de R-cuadrado ajustado para modelos DiD asociados al plebiscito nacional por tratamiento	37
3.19.	Valores de R-cuadrado ajustado para modelos DiD asociados a segunda vuelta presidencial por tratamiento	37
3.20.	Valores <i>AIC</i> para la evaluación de modelos logit multinomiales ML1 a ML5 . .	38
4.1.	Estadísticas descriptivas: características <i>mesas y comunas</i> correspondientes a la segunda vuelta presidencial de 2017	42
4.2.	Matriz de correlación de variables COVID-19 para el plebiscito nacional de 2020	43
4.3.	Matriz de correlación de variables COVID-19 para la segunda vuelta presidencial de 2021	43

4.4.	Resumen de pruebas Levene para testear por homogeneidad de varianza	43
4.5.	Resultados de estimación <i>DiD</i> para determinar significancia de <i>Fases123</i> en la proporción de votantes por grupo etario en la segunda vuelta presidencial . . .	44
4.6.	Resultados de estimación <i>DiD</i> para determinar significancia de <i>Incidencia</i> en la proporción de votantes por grupo etario en la segunda vuelta presidencial . . .	45
4.7.	Prueba de tendencias paralelas pre-tratamiento (<i>Fases12</i>)	46
4.8.	Prueba de tendencias paralelas pre-tratamiento (<i>Incidencia</i> correspondiente a 2020)	46
4.9.	Prueba de tendencias paralelas pre-tratamiento (<i>Fases123</i>)	47
4.10.	Prueba de tendencias paralelas pre-tratamiento (<i>Incidencia</i> correspondiente a 2021)	47

Índice de Ilustraciones

1.1.	Evolución de tasa nacional de casos diarios y conteo de fallecimientos debido a COVID-19 entre abril de 2020 y julio de 2022. Fuente: Elaboración propia utilizando datos de incidencia y defunciones publicados por el MinCiencia. . .	2
1.2.	Relación caso-mortalidad por país. Fuente: <i>John Hopkins Coronavirus Resource Center</i>	3
1.3.	Ocupación de camas UCI en la Región Metropolitana entre marzo de 2020 y septiembre de 2022. Fuente: DataGob Universidad Católica de Chile.	3
1.4.	Respuestas a la pregunta: <i>¿Cuánto le preocupan las siguientes situaciones sobre concurrir a votar al plebiscito?</i> Fuente: Encuesta IPSOS-Espacio Público 2020.	4
1.5.	Distribución de respuestas de la pregunta: <i>¿Cree usted que aumentarán los contagios de COVID en el país después de la realización del plebiscito?</i> Fuente: Encuesta IPSOS-Espacio Público 2020.	4
1.6.	Evolución de participación electoral entre 2017-2021 a nivel nacional. Fuente: Elaboración propia utilizando datos de resultados electorales publicados por el Servicio Electoral.	5
1.7.	Diferencial de participación nacional por grupo etario entre el plebiscito de 2020 y la segunda vuelta presidencial de 2017. Fuente: Elaboración propia utilizando datos de resultados electorales publicados por el Servicio Electoral.	6
1.8.	Diferencial de participación nacional por grupo etario entre la segunda vuelta presidencial de 2021 y 2017. Fuente: Elaboración propia utilizando datos de resultados electorales publicados por el Servicio Electoral.	6
1.9.	Composición etaria de votantes entre la segunda vuelta presidencial de 2017 y el plebiscito nacional de 2020. Fuente: Elaboración propia utilizando datos de resultados electorales publicados por el Servicio Electoral.	7
1.10.	Composición etaria de votantes en la segunda vuelta presidencial de 2017 y 2021. Fuente: Elaboración propia utilizando datos de resultados electorales publicados por el Servicio Electoral.	7
3.1.	Proceso <i>CRISP-DM</i>	14

Capítulo 1

INTRODUCCIÓN

1.1. El Sistema Electoral de Chile

El sistema electoral en Chile ha experimentado cambios significativos a lo largo de su historia. Estos incluyen el retorno a un sistema democrático de elecciones en 1988, cambios en la obligatoriedad de voto en distintos periodos y el establecimiento del proceso de elección de gobernadores regionales en 2018. En la actualidad el sistema electoral se rige por un sistema proporcional (método D'Hondt) para la elección de diputados y senadores, así como por un sistema de votación mayoritario para las elecciones presidenciales y las elecciones de consejeros regionales. Por ejemplo, si ningún candidato presidencial obtiene más del 50 % de los votos en la primera vuelta, se realiza una segunda vuelta entre los dos candidatos con mayor cantidad de votos, donde un candidato es declarado electo en el caso de obtener la mayoría simple.

Desde que comenzó a regir la Ley N° 20.568 en 2012, la inscripción de ciudadanos mayores a 17 años en el registro electoral es automática. Esta misma ley cambió la obligatoriedad del voto a un carácter voluntario, lo cual duró 10 años hasta el plebiscito nacional de 2022. Previo a esta elección, cada ciudadano concretaba el voto en una misma mesa de votación ubicado en su territorio electoral correspondiente. Los territorios electorales son divididos en zonas geográficas en varios niveles de agregación. Estos incluyen las circunscripciones senatoriales, distritos, circunscripciones provinciales y las circunscripciones electorales. La circunscripción electoral corresponde al nivel más desagregado dado que consiste en divisiones territoriales dentro de cada comuna.

A pesar de lo fundamental que es la participación en toda elección política, han habido elecciones particularmente importantes durante los últimos cinco años en Chile. El 25 de octubre de 2020, el 51 % de la población inscrita votó mediante un plebiscito nacional aprobar o rechazar la entrada a un nuevo proceso de reforma constitucional, 40 años después de la reforma anterior. Con un resultado de casi 80 % por la opción de *Apruebo*, esta elección dio paso a la elección de convencionales constituyentes en mayo de 2021, quienes tuvieron la responsabilidad de elaborar y proponer una nueva versión de la constitución. En noviembre y diciembre del mismo año, se realizaron las elecciones presidenciales, resultando en la elección de Gabriel Boric como presidente de Chile para el periodo de 2022 a 2026.

1.2. Pandemia en Chile

Es importante considerar que las elecciones mencionadas anteriormente se realizaron bajo el contexto de la pandemia de COVID-19, una época que tuvo un impacto significativo en la economía, el sistema de salud y la movilidad de los ciudadanos desde que el primer caso fue detectado el 3 de marzo de 2020 en la ciudad de Talca. La Figura 1.1 muestra las instancias de elecciones nacionales que se realizaron durante la alerta sanitaria con respecto a la evolución del virus. De acuerdo al último informe epidemiológico publicado por el Ministerio de Salud, el COVID-19 quitó un total de 57.852 vidas en Chile. En la Figura 1.2 se puede ver la posición de Chile con respecto a otros países en mortalidad post pandemia, donde Chile tuvo una tasa de caso-mortalidad del 1,2 %, es decir, por cada 100 casos de contagio hubieron 1,2 casos fallecidos, posicionando a Chile por arriba de Estados Unidos y varios países europeos. En la figura que sigue se puede ver que la tasa de ocupación de camas UCI en la Región Metropolitana estuvo en niveles críticos (sobre 85 %) a lo largo de los dos años de la alerta sanitaria.

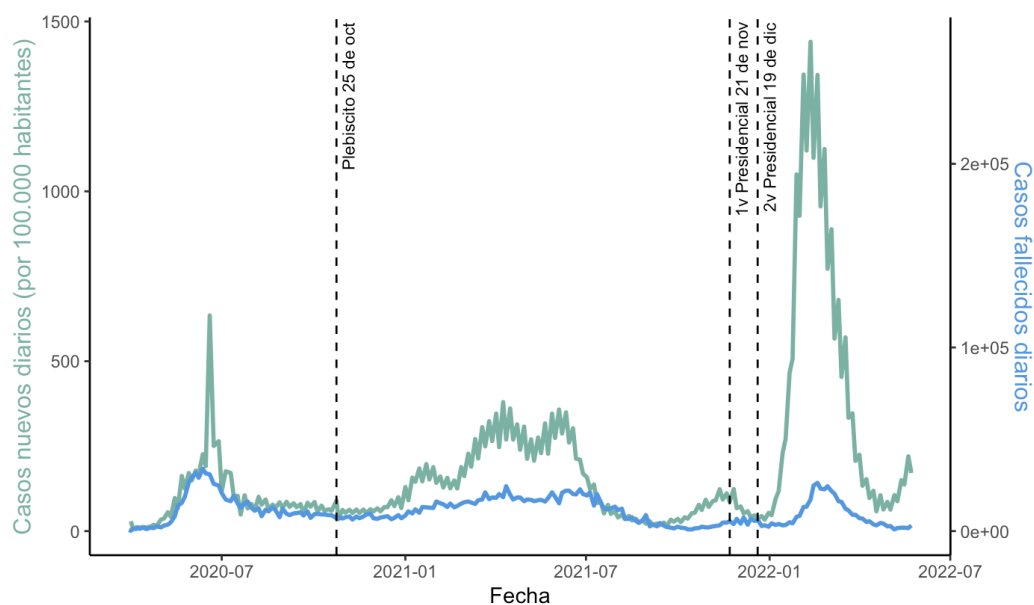


Figura 1.1: Evolución de tasa nacional de casos diarios y conteo de fallecimientos debido a COVID-19 entre abril de 2020 y julio de 2022. Fuente: Elaboración propia utilizando datos de incidencia y defunciones publicados por el MinCiencia.

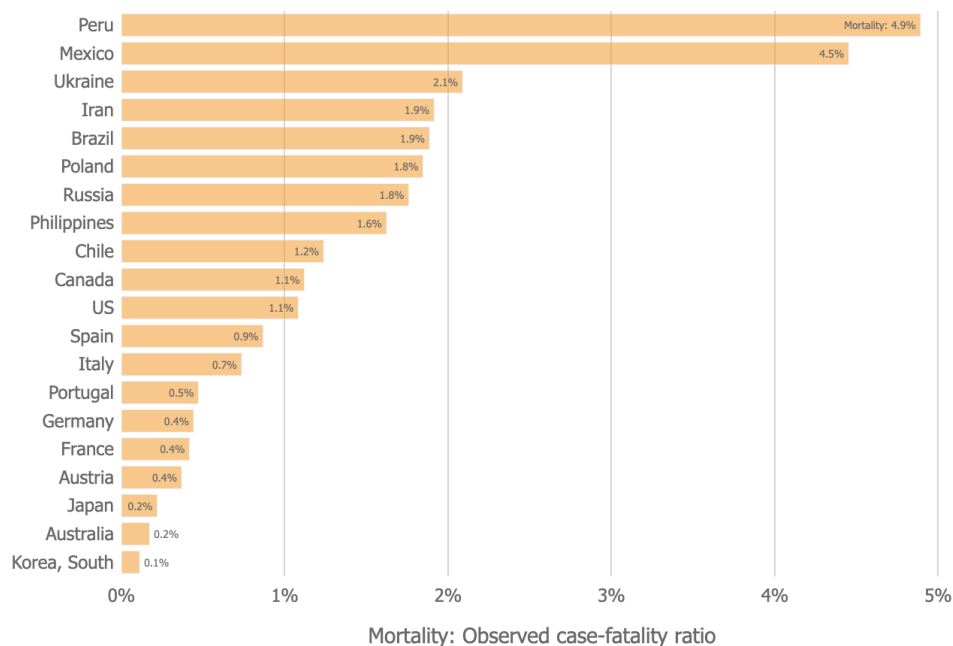


Figura 1.2: Relación caso-mortalidad por país. Fuente: *John Hopkins Coronavirus Resource Center*.

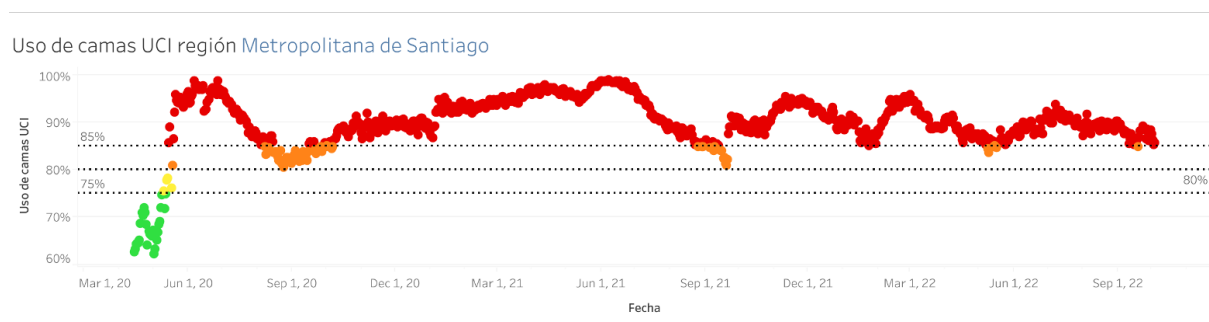


Figura 1.3: Ocupación de camas UCI en la Región Metropolitana entre marzo de 2020 y septiembre de 2022. Fuente: DataGob Universidad Católica de Chile.

En cuanto al impacto económico del virus, la tasa de desocupación alcanzó un 13,1 % en julio de 2020, subiendo 5,5 p.p. desde el año anterior (fuente: Banco Central). Como respuesta a las altas tasas de ocupación de camas UCI y la propagación del virus, el Ministerio de Salud implementó el Plan Paso a Paso el cual consistió en una estrategia gradual de desconfinamiento de 5 etapas con el fin de mitigar el virus. Esta iniciativa implicó distintas restricciones de movilidad ciudadana por distintos periodos de tiempo en cada zona geográfica. Por ejemplo, desde el inicio del Plan Paso a Paso hasta el plebiscito nacional en octubre de 2020, la zona urbana de la comuna de Arica estuvo bajo la primera fase del plan (mayores restricciones) durante todo el periodo, es decir aproximadamente 90 días, mientras que la zona rural de la misma comuna se encontraba en la tercera fase.

Sin embargo, las restricciones de confinamiento no fueron la única causa de cambios en movilidad en las ciudades, pues existió una amenaza percibida del virus que también influyó en

la disminución de movimiento. Los resultados de la Encuesta IPSOS-ESPACIO PÚBLICO publicados en agosto de 2020 mostraron que más del 52 % de los respondientes declararon tener alto nivel de preocupación de ser contagiado sobre concurrir a votar al plebiscito del 25 de octubre. Además, el 47 % creyeron que aumentarían los contagios producto de la elección.



Figura 1.4: Respuestas a la pregunta: *¿Cuánto le preocupan las siguientes situaciones sobre concurrir a votar al plebiscito?* Fuente: Encuesta IPSOS-Espacio Público 2020.

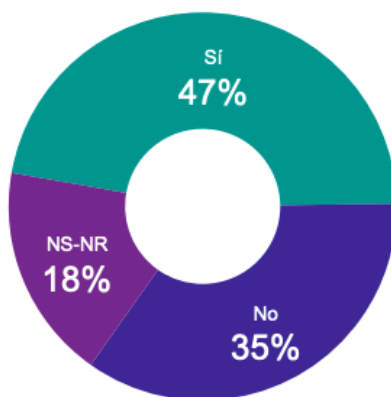


Figura 1.5: Distribución de respuestas de la pregunta: *¿Cree usted que aumentarán los contagios de COVID en el país después de la realización del plebiscito?* Fuente: Encuesta IPSOS-Espacio Público 2020.

A pesar de la amenaza real y percibida del virus, el plebiscito nacional de 2020 y la segunda vuelta presidencial de 2021 fueron las elecciones de mayor participación voluntaria histórica hasta la fecha, donde más de 50 % de los inscritos decidieron votar. Si se analizan los cambios en participación nacional para distintos rangos etarios con respecto a las elecciones presidenciales de 2017, en la Figura 1.7 se puede ver que todos los grupos etarios aumentaron su participación durante el plebiscito nacional, donde el grupo de mayor incremento correspondió al rango de votantes entre 20-24 años de edad (+46 pp). A pesar del incremento para todos los grupos en el plebiscito, se puede ver en la misma figura que los niveles de participación aumentaron a tasas menores para votantes mayores. Sin embargo, las tasas de cambio fueron más similares entre grupos durante la segunda vuelta presidencial con excepción de los

votantes mayores a 75 años. En cuanto a las diferencias en composición etaria de la población votante relativo a 2017, en las figuras 1.9 y 1.10 se puede ver que la proporción de votantes jóvenes entre 18-39 años de edad incrementó, mientras que el porcentaje de votantes mayores a 40 disminuyó para ambas elecciones de pandemia.

Dado el contexto explicado anteriormente, surge la pregunta: **¿es posible que las implicancias de la pandemia, particularmente las restricciones de movilidad y la propagación del virus, tuvieron un impacto en quiénes votaron y cómo votaron los ciudadanos durante las elecciones?** En el presente trabajo se investiga acerca la pregunta planteada comenzando con una revisión de literatura relevante al tema y el planteamiento de los objetivos y alcance del trabajo (Capítulo 1). Luego se presenta el marco teórico donde se explica la base conceptual de los métodos econométricos utilizados (Capítulo 2), dando paso al desarrollo del proceso metodológico implementado para poder cumplir con los objetivos del trabajo (Capítulo 3). Finalmente, se concluye acerca los resultados y se proponen futuros trabajos que podrían contribuir al tema de investigación (Capítulo 4).

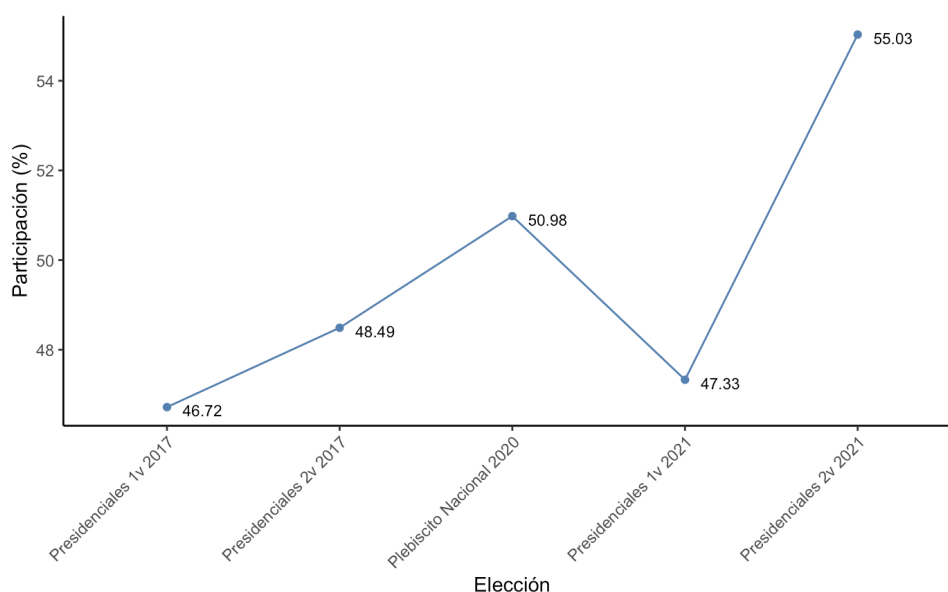


Figura 1.6: Evolución de participación electoral entre 2017-2021 a nivel nacional. Fuente: Elaboración propia utilizando datos de resultados electorales publicados por el Servicio Electoral.

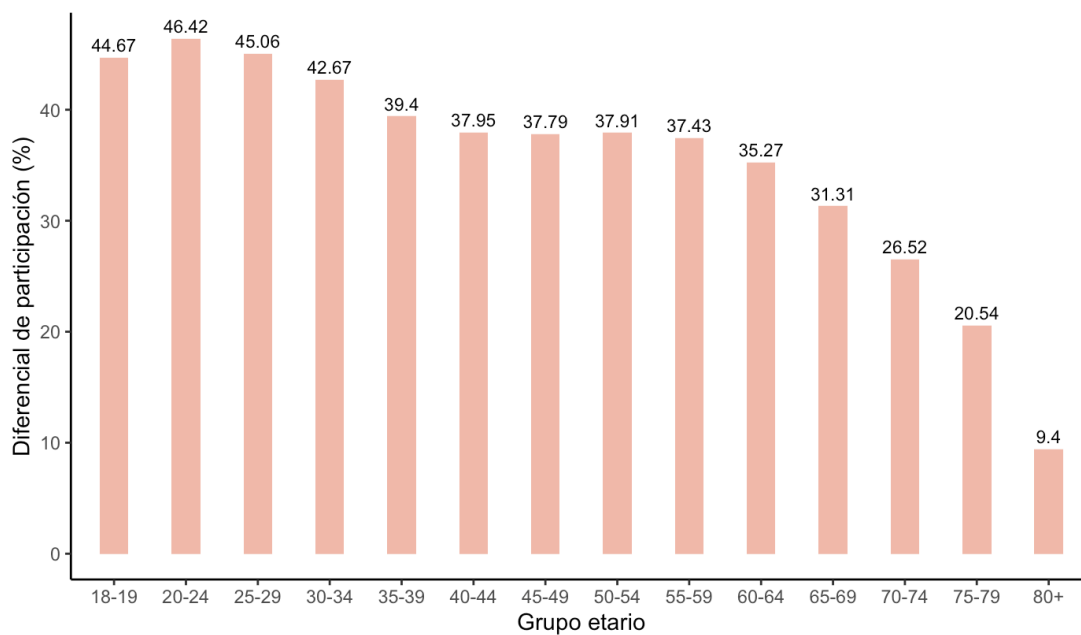


Figura 1.7: Diferencial de participación nacional por grupo etario entre el plebiscito de 2020 y la segunda vuelta presidencial de 2017. Fuente: Elaboración propia utilizando datos de resultados electorales publicados por el Servicio Electoral.

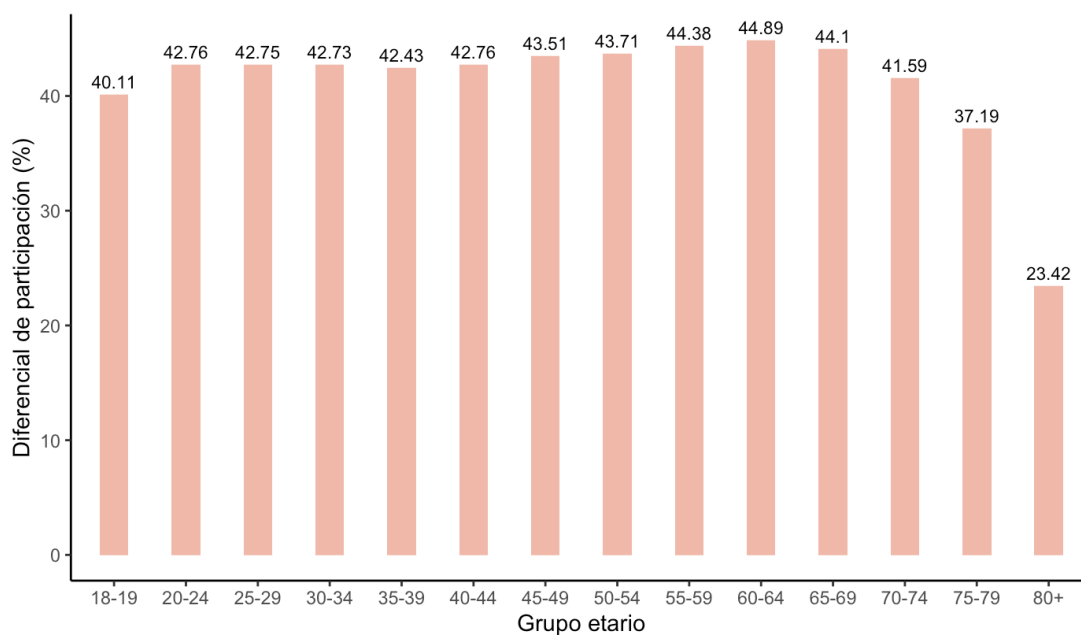


Figura 1.8: Diferencial de participación nacional por grupo etario entre la segunda vuelta presidencial de 2021 y 2017. Fuente: Elaboración propia utilizando datos de resultados electorales publicados por el Servicio Electoral.

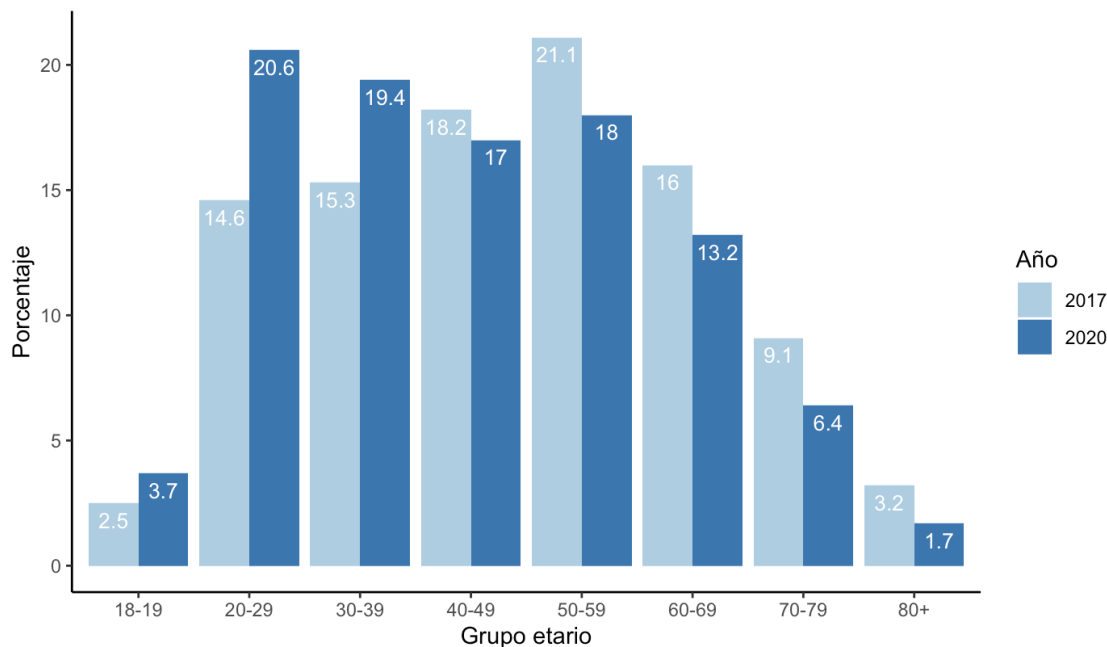


Figura 1.9: Composición etaria de votantes entre la segunda vuelta presidencial de 2017 y el plebiscito nacional de 2020. Fuente: Elaboración propia utilizando datos de resultados electorales publicados por el Servicio Electoral.

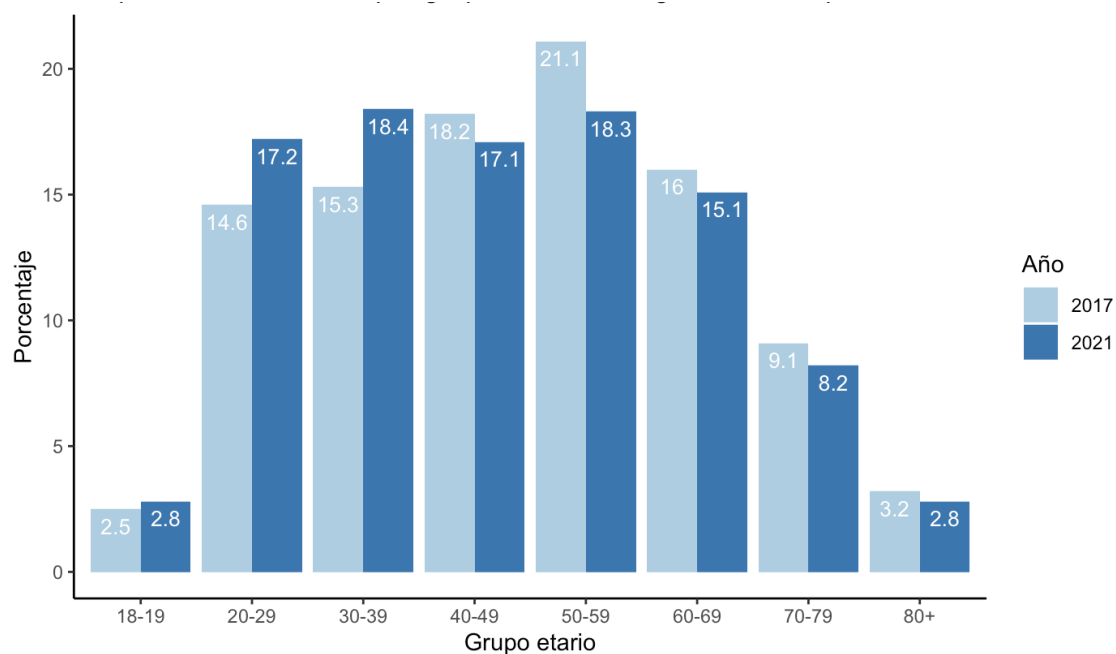


Figura 1.10: Composición etaria de votantes en la segunda vuelta presidencial de 2017 y 2021. Fuente: Elaboración propia utilizando datos de resultados electorales publicados por el Servicio Electoral.

1.3. Literatura

El impacto de COVID-19 en las elecciones políticas sigue siendo un tema reciente, sin embargo esta relación se ha investigado en varios países pues Chile no fue el único caso donde se ejecutaron elecciones políticas durante el periodo de pandemia. En cuanto a los efectos en la participación electoral, Scheller [2021] investigó acerca la relación entre la tasa de participación y la interacción entre contagios con edad en Florida, uno de los estados con la porción más alta de la población mayor en Estados Unidos. Se encontró que electores mayores en edad y en zonas con mayor tasa de contagios fueron menos probables de votar, ya sea en persona o por correo, lo cual puede ser debido a que los riesgos de infección y mortalidad fueron más elevados para la población mayor. Esto también es consistente con los resultados de Mondschein et al. [2022], quienes encontraron que el riesgo de infección fue una factor importante en la participación electoral durante el plebiscito nacional de 2020.

Estados Unidos es un caso notable donde se observó una fuerte polarización entre el Partido Demócrata y Partido Republicano con respecto a la amenaza del virus y las medidas a implementar, resultando en la posibilidad de que el COVID-19 incide en cómo votaran los ciudadanos. Warshaw et al. [2020] mostraron que el ex-presidente Donald Trump fue menos probable de ser reelecto en 2020 en estados con mayores niveles de fallecimientos asociados a COVID-19, posiblemente debido al descontento de algunos votantes en los estados mayormente impactados con respecto a su falta de iniciativa en la implementación de medidas preventivas a nivel federal. Sin embargo, Karwowski et al. [2020] encontró que la amenaza de COVID-19 durante las elecciones presidenciales en Estados Unidos y Polonia aumentó la preferencia por candidatos conservadores.

Giommoni and Loumeau [2020] estudiaron el efecto de las restricciones de confinamiento en los resultados de las elecciones municipales de Francia. Encontraron que las zonas municipales con restricciones más duras votaron distinto a las zonas con restricciones más leves de manera significativa, donde el primer caso fue asociado con un mayor apoyo por el Partido Verde. También, mostraron que mayores restricciones fueron asociados con una mayor participación, posiblemente debido a que el tiempo en confinamiento incentivaron a los votantes a movilizarse. Otro estudio de las mismas elecciones municipales en Francia por Adam-Troian et al. [2023] muestra que el sentido de amenaza de COVID-19, definido a través del volumen de búsqueda en Google, fue asociado a un incremento de votos por partidos conservadores.

1.4. Objetivos

Objetivo general

A partir de la pregunta de investigación planteada anteriormente, se define el siguiente objetivo general: analizar el efecto de factores observables asociados a COVID-19 en la composición demográfica de los votantes y en las preferencias de voto durante las elecciones políticas de pandemia.

Objetivos específicos

- Analizar las tendencias de preferencia y composición demográfica en elecciones antes y durante la pandemia de COVID-19.
- Proponer variables observables explicativas asociadas a COVID-19 y analizar sus efectos en la composición demográfica de los votantes.
- Analizar la relación entre las variables propuestas con las preferencias de voto.

1.5. Alcance

El caso de estudio se enfoca en realizar un análisis descriptivo de los efectos de las variables de interés en las siguientes elecciones políticas que ocurrieron en Chile durante el periodo de la pandemia de COVID-19: el plebiscito nacional de 2020 y las elecciones presidenciales de 2021, donde se utilizan los resultados de las elecciones presidenciales de 2017 como periodo de comparación. El análisis se centra en la población de ciudadanos que sufragaron, considerando sólo votos emitidos en territorio nacional.

Capítulo 2

MARCO TEÓRICO

2.1. Métodos

2.1.1. Regresión lineal multiple

La regresión multiple es uno de los métodos más utilizados para modelar las relaciones entre variables. En su forma lineal, queda de la siguiente manera:

$$\begin{aligned} Y &= X\beta + \epsilon \\ &= \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + \epsilon_i \quad \forall i = 1, \dots, n \end{aligned} \tag{2.1}$$

Se tiene el vector Y de las variables de resultado de cada observación, la matriz de variables explicativas X , el vector de coeficientes β y el vector de residuos ϵ . Una vez especificado el modelo, se desea estimar los parámetros β_1, \dots, β_k , los cuales representan los efectos promedios de los regresores en la variable dependiente. La estimación se realiza a través del método de mínimos cuadrados ordinarios, lo cual minimiza la suma de los residuos cuadrados. Es importante que se cumplen los supuestos del método *OLS* para una estimación libre de sesgo:

- i) No hay correlación entre las variables explicativas y los errores, i.e. endogeneidad.
- ii) Los errores tienen una varianza fija entre observaciones y son independientes, es decir, distribuyen *i.i.d.*
- iii) No existe multicolinealidad perfecta, i.e. una fuerte o perfecta correlación entre variables independientes.

2.1.2. Regresión logística multiple o Logit binomial

El resultado de una elección con dos posibles opciones de resultado se puede representar de forma binaria o binomial. Por ejemplo, el resultado de la segunda vuelta presidencial anterior puede definirse de las siguientes formas:

- a) $Y = 1$ si gana Boric; $Y = 0$ si gana Kast (binaria o binomial con $n = 1$)
- b) $Y =$ cantidad de votos por Kast (binomial con $n =$ votos totales)
- c) $Y =$ cantidad de votos por Boric (binomial con $n =$ votos totales)

En este caso se modela la relación de la variable de resultado con un conjunto de p variables explicativas mediante la regresión logística binomial.

$$\log\left(\frac{\pi_i}{1 - \pi_i}\right) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ip} \quad (2.2)$$

$$\text{donde } \pi_i = \frac{1}{1 + \exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ip})}$$

El método de estimación más común es el estimador de máximo verosimilitud (*MLE*) el cual encuentra los parámetros que maximizan la probabilidad de observar los datos dados. Al igual que la regresión lineal *OLS*, la regresión logística asume errores *i.i.d* y la ausencia de endogeneidad y multicolinealidad.

2.1.3. Diferencia-en-Diferencias (DiD)

Diferencia-en-diferencias (DiD) es una técnica econometrica tipicamente utilizada para analisis de impacto, pues permite medir el efecto de una intervención. Este efecto se define mediante la siguiente ecuación:

$$\begin{aligned} \text{DiD} &= \mathbb{E}[Y_{it}(1, T) - Y_{it}(0, T)|D = 1] - \mathbb{E}[Y_{it}(1, C) - Y_{it}(0, C)|D = 0] \\ &= \mathbb{E}[\Delta_T|D = 1] - \mathbb{E}[\Delta_C|D = 0] \text{ donde} \end{aligned} \quad (2.3)$$

- $Y_{it}(1, T)$ representa el resultado potencial del individuo i en el tiempo t perteneciente al grupo de tratamiento al recibir el tratamiento
- $Y_{it}(0, T)$ representa el resultado potencial del individuo i en el tiempo t perteneciente al grupo de tratamiento sin recibir el tratamiento
- $Y_{it}(1, C)$ es el resultado potencial de i en t perteneciente al grupo de control al recibir el tratamiento (contrafactual)
- $Y_{it}(0, C)$ es el resultado potencial de i en t perteneciente al grupo de control sin recibir el tratamiento
- D es el indicador de asignación de tratamiento, donde $D = 1$ corresponde al grupo de tratados y $D = 0$ al grupo de control
- $\Delta_T = Y_{it}(1, T) - Y_{it}(0, T)$ representa la diferencia potencial de resultados para el grupo de tratamiento
- $\Delta_C = Y_{it}(1, C) - Y_{it}(0, C)$ representa la diferencia potencial de resultados para el grupo de control

Algunas aplicaciones incluyen la evaluación de campañas de marketing o de políticas públicas, ya sea en salud o educación. Al aplicar el modelo es importante que se cumplan los supuestos necesarios para garantizar la validez de las estimaciones. El supuesto más importante para el

análisis causal es el de tendencias paralelas, donde se asume que la diferencia en las medias de las variables dependientes entre el grupo tratado y el grupo de control permanecen constantes en ausencia del tratamiento. Otro supuesto es que la asignación del tratamiento debe ser aleatoria y exógena, es decir no correlacionada con factores no observables. En su forma lineal, el modelo queda de la siguiente manera, donde $Treatment_i$ es el identificador del grupo de tratamiento, $Post_t$ identifica el periodo después de la intervención y X_{it} representa un vector de covariables:

$$Y_{it} = \beta_0 + \beta_1 Treatment_i + \beta_2 Post_t + \beta_3 (Treatment_i \times Post_t) + \gamma X_{it} + \epsilon_{it} \quad (2.4)$$

2.1.4. Logit multinomial

Cuando un tomador de decisión como por ejemplo un votante debe elegir entre más de dos alternativas o candidatos, se puede modelar este comportamiento mediante modelos de elección discreta. La estructura del modelo es de la siguiente forma: un individuo n debe elegir entre J alternativas, escogiendo la alternativa j que le entregue mayor utilidad. La definición del conjunto de alternativas debe cumplir con tres características: (1) la elección debe ser mutuamente excluyente, es decir, el tomador de decisión debe elegir una sola alternativa, (2) el conjunto de alternativas debe ser exhaustiva, es decir, incluye todas las posibles opciones, (3) y el conjunto debe ser finita.

La utilidad de cada individuo n es desconocida por el investigador, por lo que se descompone por un componente observable de la utilidad y un componente no-observable (estocástico) V_{nj} y la utilidad no-observable ε_{nj} . La utilidad observable $V_{nj} = x'_{nj}\beta$ representa los factores observados del investigador, ya sean atributos de los individuos o de las alternativas.

$$U_{nj} = V_{nj} + \varepsilon_{nj} \quad (2.5)$$

A partir de la especificación de la utilidad, se puede definir las probabilidades de elección para cada individuo.

$$P_{ni} = Prob(U_{ni} > U_{nj} \forall j \neq i) \quad (2.6)$$

En el caso del modelo logit, una extensión del logit binomial, la probabilidad de elección tiene la siguiente forma cerrada:

$$P_{ni} = \frac{e^{\beta' x_{ni}}}{\sum_j e^{\beta' x_{nj}}} \quad (2.7)$$

Los supuestos de este modelo son los siguientes:

- La utilidad no observada distribuye independientemente e idénticamente valor extremo
- Independencia de alternativas irrelevantes (IIA), es decir, la relación de *odds* (razón de probabilidades) entre dos alternativas de la variable dependiente no debe cambiar al introducir una tercera opción

Capítulo 3

DESARROLLO METODOLÓGICO

3.1. Metodología *CRISP-DM*

La metodología *CRISP-DM* o *Cross-Industry Standard Process for Data Mining* es una de las más utilizadas en varios sectores para proyectos de minería de datos dado su robustez y versatilidad. Como se puede ver en la siguiente figura, el proceso tradicional consiste en seis etapas:

1. Entendimiento del negocio (*Business Understanding*) - ¿Cuáles son los objetivos del negocio? ¿Cuáles son las metas técnicas en términos de minería de datos?
2. Entendimiento de los datos (*Data Understanding*) - ¿Qué tipo de información se requiere? ¿Cómo es la calidad de los datos disponibles?
3. Preparación de datos (*Data Preparation*) - ¿Cómo se deben transformar y formatear los datos?
4. Modelamiento (*Modeling*) - ¿Cuáles métodos de modelamiento se utilizarán?
5. Evaluación (*Evaluation*) - ¿Cuáles son los criterios de evaluación de los modelos? ¿Los resultados cumplen con los objetivos del proyecto?
6. Despliegue (*Deployment*) - ¿Cuál es el plan de monitoreo y mantenimiento para que los resultados del modelo puedan ser accesibles?

En las siguientes secciones se explica en detalle el desarrollo de cada etapa del proceso *CRISP-DM*, el cual fue modificado de acuerdo a los objetivos y alcance del presente trabajo, terminando con 5 etapas. El proceso se inicia a través del entendimiento del problema de investigación (análogo al entendimiento del negocio) y se finaliza con una evaluación de los métodos aplicados, sin requerir de una etapa de despliegue.

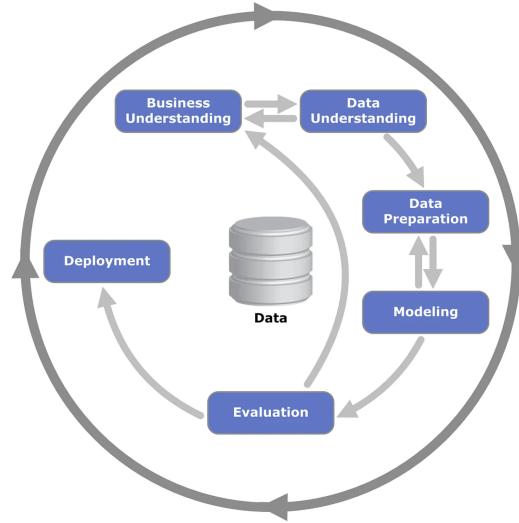


Figura 3.1: Proceso *CRISP-DM*

3.2. Entendimiento del problema

El entendimiento o planteamiento del problema consiste en definir las hipótesis de investigación. En el presente trabajo se desean analizar dos posibles efectos de la crisis sanitaria en las elecciones. El primero corresponde a un cambio en el perfil demográfico de la población votante debido al surgimiento de nuevos votantes y/o el retiro de votantes anteriores. El segundo efecto es un cambio de preferencia de voto por parte de votantes anteriores que participaron en las elecciones de pandemia. En cuanto a los factores observables asociados al COVID-19 a analizar, se estudian los efectos de la incidencia del virus, es decir la tasa de casos nuevos para un periodo de tiempo dado, y la duración de las fases más estrictas del Plan Paso a Paso, planteando las siguientes hipótesis:

- H1.** *La incidencia de COVID-19 y el tiempo en confinamiento afectaron significativamente la composición etaria de la población votante en el plebiscito nacional y las elecciones presidenciales. Particularmente, se plantea que:*
- a) *Periodos prolongados de confinamiento fueron asociados a una mayor proporción de votantes jóvenes.*
 - b) *Altos niveles de incidencia fueron asociados a una menor proporción de votantes mayores en edad.*
- H2.** *La incidencia de COVID-19 y el tiempo en confinamiento incidieron en cambios de preferencia de voto por parte de votantes ubicados en zonas con altos niveles de incidencia y bajo mayor tiempo en confinamiento.*

3.3. Entendimiento de los datos

3.3.1. Descripción de datos

A continuación se explica en detalle cuáles fueron los datos consolidados para el análisis. Los resultados de cada elección política, es decir el conteo de votos por cada alternativa o candidato, son publicados a nivel de mesa de votación por el Servicio Electoral (Servel). El Servel también publica la cantidad de votantes por género y rango etario, junto al total de inscritos, a nivel de mesa. Cabe destacar que la distribución de votantes por género y rango etario entre los candidatos/alternativas de voto no es conocida dado que el voto es anónimo. Finalmente, se puede adquirir la cantidad de inscritos por rango etario y género a nivel de comuna (pero no a nivel de mesa) a través del padrón electoral, también publicado cada año por el Servel.

En cuanto a los datos relacionados a la pandemia de COVID-19, como casos confirmados y días en cada fase del Plan Paso a Paso, estos fueron publicados a nivel comunal por el Ministerio de Ciencia¹. En la siguiente tabla se pueden ver las variables definidas a partir de los datos mencionados anteriormente junto a sus descripciones respectivas.

Tabla 3.1: Definición de variables

Variable	Definición
$VotantesM_N_{it}$	Proporción de votantes en la mesa i entre M y N años de edad en la elección t
$Mujer_{it}$	Proporción de votantes mujeres en la mesa i en la elección t
$Participacion_{it}$	Proporción de inscritos en la mesa i que votaron
$PromInscMesa_{it}$	Promedio de inscritos por mesa en el local donde se encuentra i en la elección t
$InscritosM_N_{it}$	Proporción de inscritos entre M y N años de edad en la comuna donde pertenece la mesa i en la elección t
$Fases12_{it}$	Días acumulados en Fases 1 y 2 en la comuna donde pertenece la mesa i durante los últimos tres meses previos a t
$Fases123_{it}$	Días acumulados en Fases 1, 2 y 3 en la comuna donde pertenece la mesa i durante los últimos tres meses previos a t
$Incidencia_{it}$	Numero de casos por 100.000 hab. desde los últimos tres meses previos a t en la comuna donde pertenece la mesa i
$Dens_{it}$	Densidad poblacional (habitantes por km ²) en la comuna donde pertenece la mesa i
C_i	Proporción de votos por el candidato C en la mesa i en la primera vuelta presidencial de 2021
$Choice_i$	El candidato con mayor proporción de votos en la mesa i en la primera vuelta presidencial de 2021

¹ Estos datos fueron actualizados cada 1-3 días desde el inicio de su registro hasta el fin de la altera sanitaria.

3.3.2. Análisis exploratorio de datos

A continuación se realiza un análisis exploratorio de datos, comenzando con las estadísticas descriptivas de las variables definidas anteriormente. La Tabla 3.2 resume los valores de las variables correspondientes al plebiscito nacional de 2020 a nivel de mesa y comuna, mientras que las tablas 3.3 y 3.4 corresponden a valores asociados a las elecciones presidenciales. Las estadísticas descriptivas correspondientes a la elección de referencia (segunda vuelta presidencial de 2017) se pueden encontrar en el Anexo A.1. Según la Tabla 3.2, en promedio, el 24 % de los votantes en cada mesa fueron electores entre 18 y 29 años de edad, subiendo 5 p.p. desde 2017, mientras que el promedio de la proporción de inscritos del mismo grupo subió de 3 % a 20 % a nivel comunal. Por otro lado, la proporción de votantes mayores a 70 años fue en promedio 9 %, bajando 3 p.p. desde 2017 a pesar de que el promedio de inscritos mayores a 70 disminuyó solo 1 p.p. a nivel comunal (de 13 % a 12 %).

En cuanto a los factores relacionados al COVID-19, el promedio de incidencia fluctuó entre las tres elecciones de pandemia, bajando de aproximadamente 1.000 casos a 600 casos por 100.000 habitantes entre el plebiscito y la primera vuelta presidencial, y luego volviendo a subir a 790 casos por 100.000 habitantes en la segunda vuelta. La cantidad de días en las primeras dos fases del Plan Paso a Paso varió entre 0 a 90 días desde los últimos tres meses previos al plebiscito nacional, donde comunas como Valparaíso, Antofagasta y Puente Alto estuvieron bajo las dos fases durante todo el periodo mientras que comunas como Los Ángeles y Puerto Varas no estuvieron ningún día bajo las primeras dos fases.

Tabla 3.2: Estadísticas descriptivas: características mesas y comunas correspondientes al **plebiscito nacional de 2020**

Statistic	N	Mean	St. Dev.	Min	Max
<i>Participacion</i>	43,574	0.51	0.11	0.003	0.91
<i>Mujer</i>	43,574	0.53	0.20	0.005	1.00
<i>Hombre</i>	43,574	0.47	0.20	0.00	1.00
<i>Votantes18_29</i>	43,574	0.24	0.22	0.0004	1.00
<i>Votantes30_39</i>	43,574	0.19	0.18	0.0004	0.96
<i>Votantes40_49</i>	43,574	0.17	0.14	0.0004	0.95
<i>Votantes50_59</i>	43,574	0.19	0.10	0.0004	0.83
<i>Votantes60_69</i>	43,574	0.14	0.08	0.0003	0.76
<i>Votantes+70</i>	43,574	0.09	0.07	0.0003	0.86
<i>PromInscMesa</i>	43,574	324.14	30.95	25.00	348.95
<i>Incidencia</i>	342	1,019.19	909.31	0.00	7,777.78
<i>Fases12</i>	342	23.87	32.38	0	90
<i>Fases34</i>	342	66.13	32.38	0	90
<i>Inscritos18_29</i>	342	0.20	0.03	0.05	0.33
<i>Inscritos30_39</i>	342	0.18	0.02	0.09	0.26
<i>Inscritos40_49</i>	342	0.17	0.02	0.10	0.39
<i>Inscritos50_59</i>	342	0.18	0.02	0.09	0.24
<i>Inscritos60_69</i>	342	0.13	0.01	0.07	0.18
<i>Inscritos+70</i>	342	0.13	0.04	0.03	0.57

Tabla 3.3: Estadísticas descriptivas: resultados **primera vuelta presidencial de 2021** y características a nivel de mesa y comuna

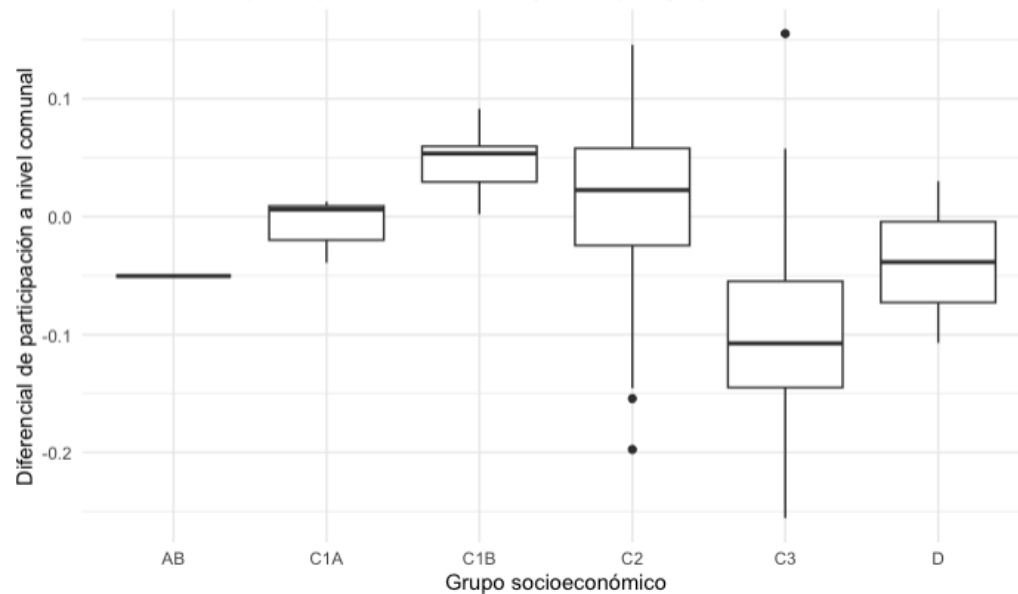
Statistic	N	Mean	St. Dev.	Min	Max
<i>Participacion</i>	45,117	0.47	0.10	0.01	0.91
<i>Mujer</i>	45,117	0.53	0.20	0.01	1.00
<i>Hombre</i>	45,117	0.47	0.20	0.03	1.00
<i>Votantes18_29</i>	45,117	0.20	0.19	0.00	1.00
<i>Votantes30_39</i>	45,117	0.18	0.18	0.00	0.97
<i>Votantes40_49</i>	45,117	0.17	0.14	0.00	0.94
<i>Votantes50_59</i>	45,117	0.19	0.09	0.00	0.84
<i>Votantes60_69</i>	45,117	0.16	0.09	0.00	0.77
<i>Votantes+70</i>	45,117	0.11	0.08	0.00	0.86
<i>PromInscMesa</i>	45,117	320.32	31.51	20.64	349.54
<i>Boric</i>	45,117	0.25	0.09	0.00	0.69
<i>Kast</i>	45,117	0.27	0.10	0.00	0.73
<i>Parisi</i>	45,117	0.13	0.09	0.00	0.69
<i>Provoste</i>	45,117	0.12	0.05	0.00	0.65
<i>Sichel</i>	45,117	0.12	0.06	0.00	0.51
<i>Ominami</i>	45,117	0.08	0.04	0.00	0.50
<i>Artes</i>	45,117	0.01	0.01	0.00	0.16
<i>Incidencia</i>	342	609.85	460.15	0.00	5,310.05
<i>Fases123</i>	342	15.63	16.34	0	76
<i>Fases45</i>	342	77.37	16.34	17	93
<i>Inscritos18_29</i>	342	0.20	0.03	0.03	0.33
<i>Inscritos30_39</i>	342	0.18	0.02	0.10	0.27
<i>Inscritos40_49</i>	342	0.17	0.02	0.10	0.37
<i>Inscritos50_59</i>	342	0.18	0.02	0.09	0.25
<i>Inscritos60_69</i>	342	0.13	0.01	0.07	0.18
<i>Inscritos+70</i>	342	0.13	0.04	0.03	0.55

Tabla 3.4: Estadísticas descriptivas: características mesas y comunas correspondientes a la **segunda vuelta presidencial de 2021**

Statistic	N	Mean	St. Dev.	Min	Max
<i>Participacion</i>	45,138	0.56	0.10	0.01	0.92
<i>Mujer</i>	45,138	0.53	0.20	0.04	0.97
<i>Hombre</i>	45,138	0.47	0.20	0.03	0.96
<i>Votantes18_29</i>	45,138	0.20	0.19	0.0003	1.00
<i>Votantes30_39</i>	45,138	0.18	0.18	0.0003	0.95
<i>Votantes40_49</i>	45,138	0.17	0.14	0.0004	0.92
<i>Votantes50_59</i>	45,138	0.18	0.09	0.0004	0.88
<i>Votantes60_69</i>	45,138	0.15	0.09	0.0003	0.77
<i>Votantes+70</i>	45,138	0.11	0.08	0.0003	0.84
<i>PromInscMesa</i>	45,138	320.32	31.52	20.64	349.54
<i>Incidencia</i>	342	791.01	671.41	0.00	8,710.80
<i>Fases123</i>	342	23.83	19.78	0	77
<i>Fases45</i>	342	68.17	19.78	15	92
<i>Inscritos18_29</i>	342	0.20	0.03	0.03	0.33
<i>Inscritos30_39</i>	342	0.18	0.02	0.10	0.27
<i>Inscritos40_49</i>	342	0.17	0.02	0.10	0.37
<i>Inscritos50_59</i>	342	0.18	0.02	0.09	0.25
<i>Inscritos60_69</i>	342	0.13	0.01	0.07	0.18
<i>Inscritos+70</i>	342	0.13	0.04	0.03	0.55

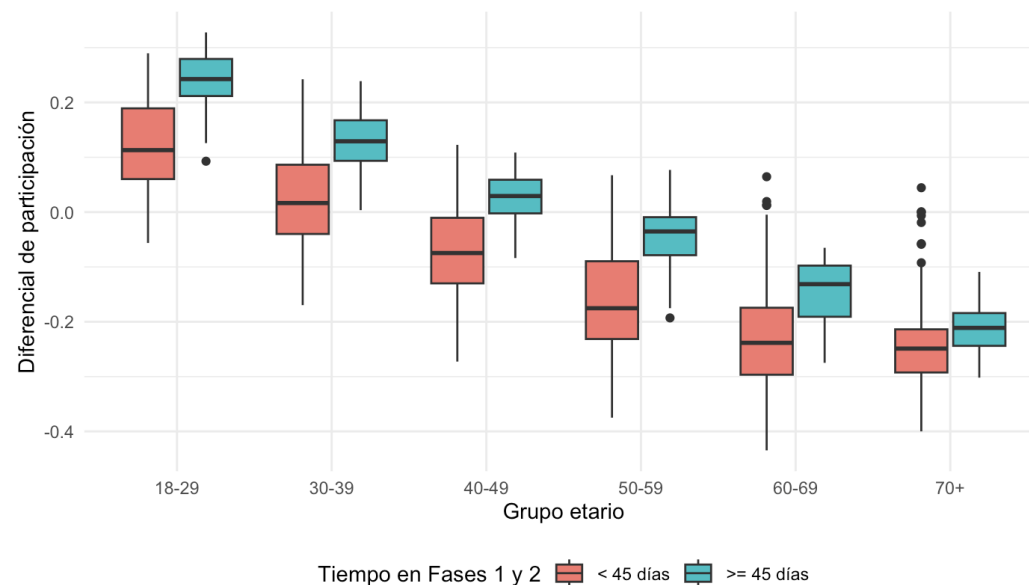
Para explorar cambios entre las elecciones estudiadas y la última elección nacional pre-pandemia se visualiza primero el diferencial de participación en las comunas agrupadas por grupo socioeconómico, donde un valor de -0.1 indica una disminución de 10 pp. En la Figura 3.2 se ve que, en promedio, los grupos extremos (AB,C3 y D) bajaron su participación mientras que los grupos intermedios (C1B,C2) aumentaron o mantuvieron su participación con respecto a la segunda vuelta presidencial de 2017. Luego, en las figuras 3.3 y 3.4 se grafican los cambios en participación a nivel de comuna por rango etario y tiempo en confinamiento. A pesar de que los niveles de participación aumentaron a nivel nacional para todos los grupos etarios, al analizar los cambios de manera desagregada se puede ver que votantes mayores en edad bajaron su participación en promedio a nivel comunal en ambas elecciones. También se puede observar que aquellas comunas que estuvieron más tiempo (≥ 45 días) en las fases más estrictas del Plan Paso a Paso tuvieron mayores niveles de participación en comparación a las demás comunas.

Figura 3.2: Diferencial de participación de comunas por grupo socioeconómico entre el plebiscito de entrada (2020) y la elección presidencial anterior (2017)



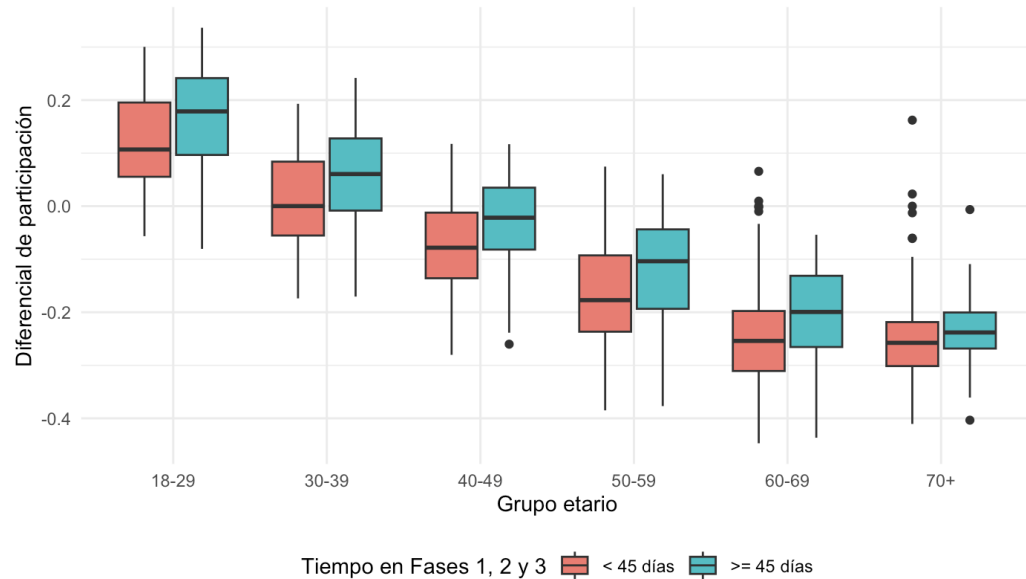
Fuente: Elaboración propia utilizando datos de participación publicados por el SERVEL e ingresos promedios por hogar obtenidos de la Encuesta CASEN de 2017.

Figura 3.3: Cambio en participación a nivel comunal por grupo etario entre el plebiscito de entrada (2020) y la elección presidencial anterior (2017)



Fuente: Elaboración propia utilizando datos de distribución de votantes publicados por el SERVEL.

Figura 3.4: Cambio en participación a nivel comunal por grupo etario entre la segunda vuelta presidencial de 2021 y 2017

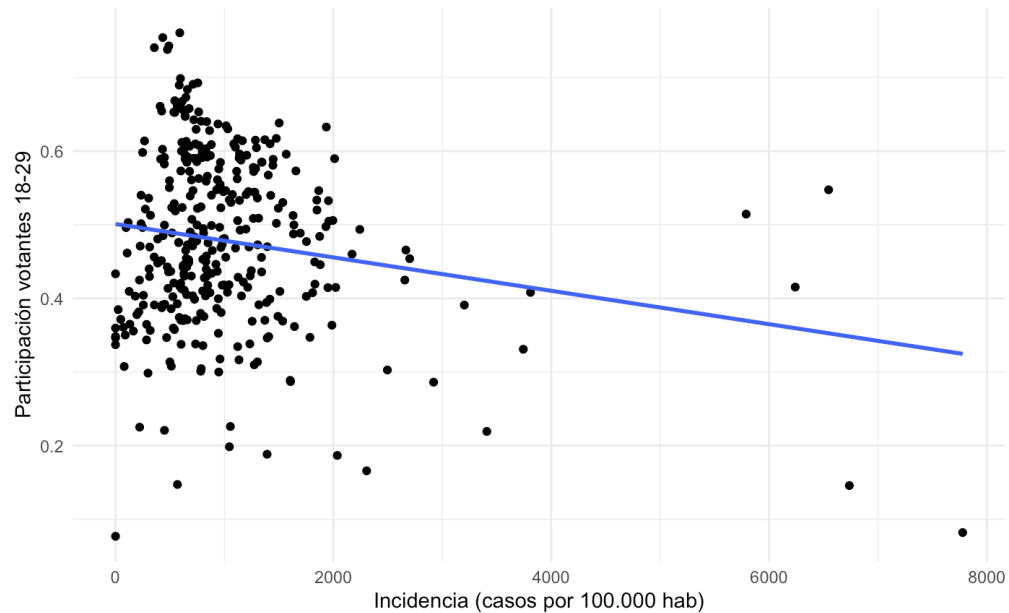


Fuente: Elaboración propia utilizando datos de distribución de votantes publicados por el SERVEL.

Al analizar los niveles de participación durante el plebiscito con respecto a los niveles de incidencia, en las figuras 3.5 a 3.7 se pueden ver que hubieron relaciones negativas entre las dos variables para los tres rangos etarios graficados (18-29, 50-59, +70), indicando menos participación en comunas con mayor incidencia. Sin embargo, a contrario de lo que se esperaba, no se observan pendientes menores para los votantes mayores. Cabe destacar que estos gráficos no toman en cuenta los niveles de participación previos a la pandemia, por lo que se procede a visualizar el diferencial de participación con respecto a la incidencia de COVID-19 en las figuras 3.8 a 3.10. En estos gráficos se observan pendientes horizontales para los votantes entre 18-29 y 70+, y un pendiente ligeramente positivo para electores entre 50-59 años. Sin embargo, no se ve una tendencia clara entre las dos variables por lo que no se puede inferir una relación positiva o negativa entre la incidencia y el diferencial de participación con respecto a 2017.

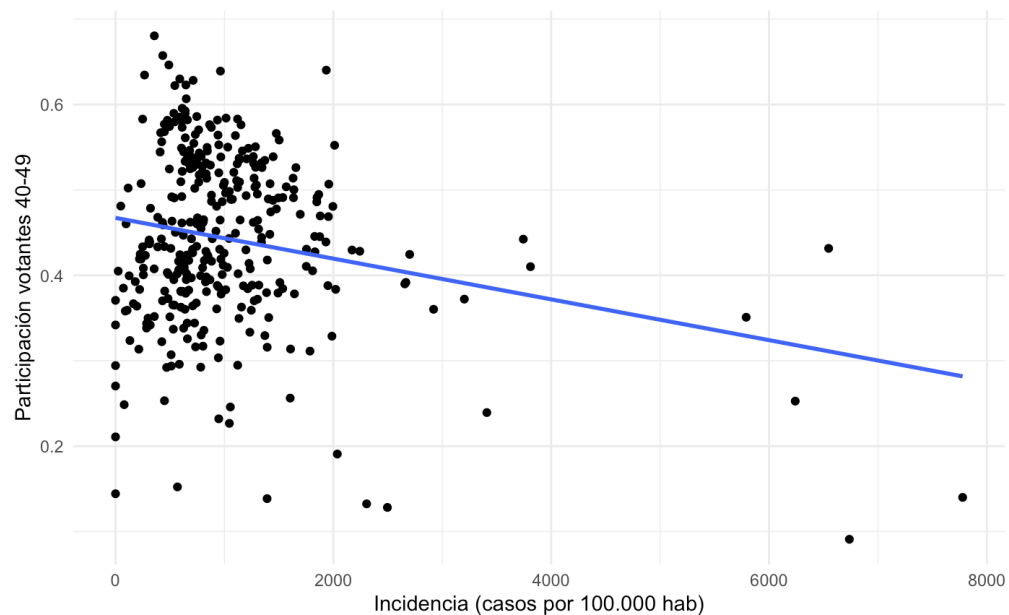
En las tablas que siguen, se resumen los resultados de varias pruebas estadísticas realizadas para determinar si hubieron diferencias significativas en los cambios en participación a nivel de comuna para distintas agrupaciones. Primero, se ejecuta una prueba Kruskal-Wallis para testear por diferencias entre los grupos etarios. Como se puede ver en la Tabla 3.5, el p-valor menor a 0.05 indica que existieron diferencias significativas entre los seis rangos etarios en términos de sus cambios en participación entre 2017 y 2020. Luego, se aplican las pruebas *t* de *Welch* y *Wilcoxon Rank Sum* para evaluar la existencia de diferencias significativas entre: (1) zonas con alta y baja incidencia, y (2) comunas que estuvieron por sobre o debajo de la mediana de días en las fases 1 y 2. Los resultados mostrados en las tablas 3.6 a 3.9 indican que no existieron diferencias significativas en los cambios en participación de cada rango etario entre zonas con alta y baja incidencia, con excepción de votantes mayores a 70 años. Sin embargo, sí existieron diferencias entre las comunas agrupadas por tiempo en las fases 1 y 2 para los primeros cinco rangos etarios.

Figura 3.5: Relación entre tasa de participación a nivel comunal y tasa de incidencia para votantes entre 18-29 años de edad en el plebiscito nacional de 2020



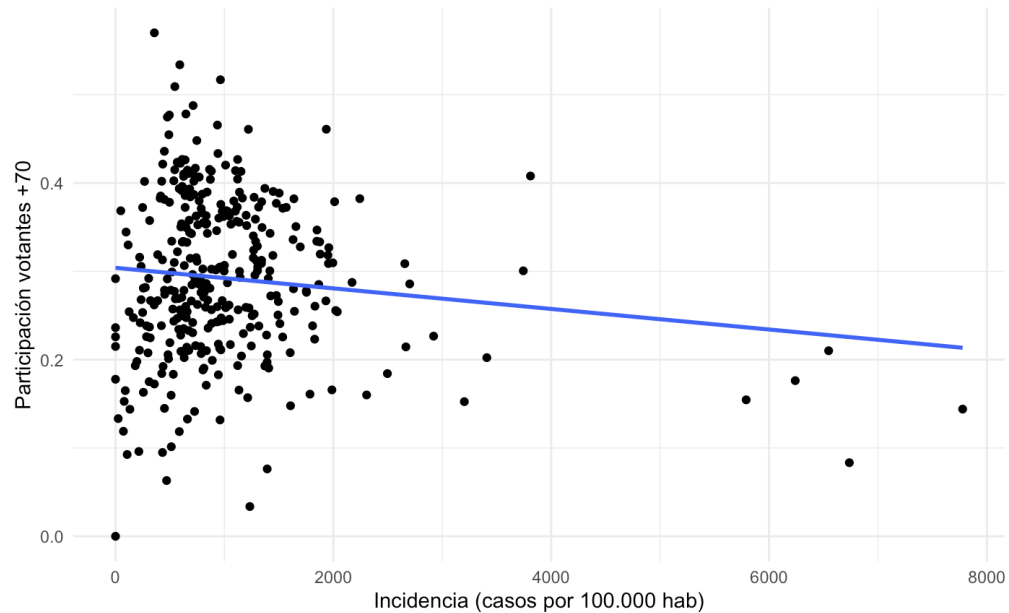
Fuente: Elaboración propia utilizando datos de distribución de votantes publicados por el SERVEL.

Figura 3.6: Relación entre tasa de participación a nivel comunal y tasa de incidencia para votantes entre 50-59 años de edad en el plebiscito nacional de 2020



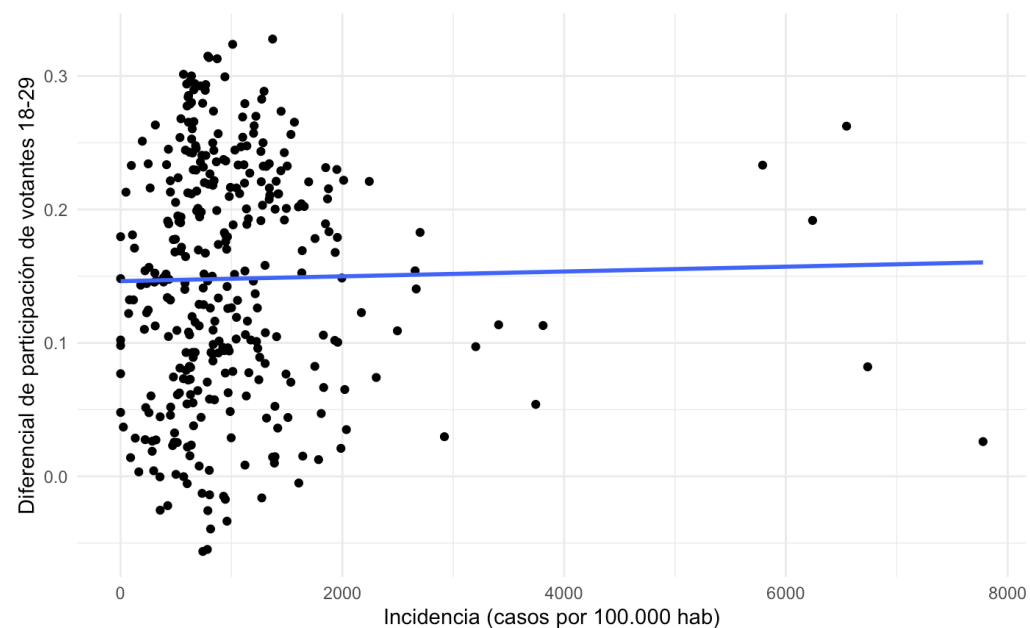
Fuente: Elaboración propia utilizando datos de distribución de votantes publicados por el SERVEL.

Figura 3.7: Relación entre tasa de participación a nivel comunal y tasa de incidencia para votantes entre 50-59 años de edad en el plebiscito nacional de 2020



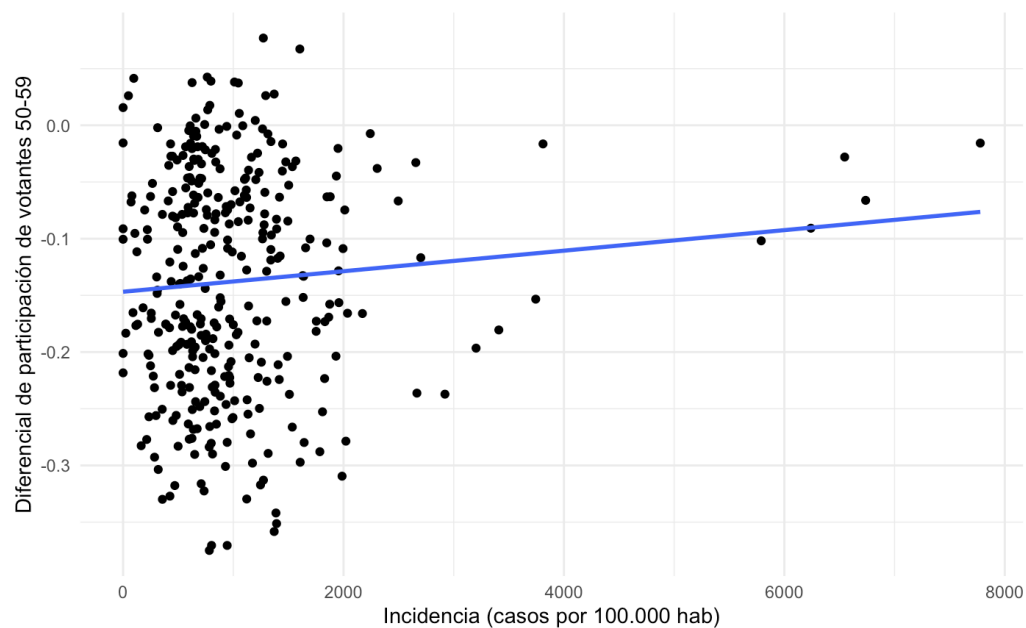
Fuente: Elaboración propia utilizando datos de distribución de votantes publicados por el SERVEL.

Figura 3.8: Relación entre diferencial de participación a nivel comunal y tasa de incidencia para votantes entre 18-29 años de edad en el plebiscito nacional de 2020



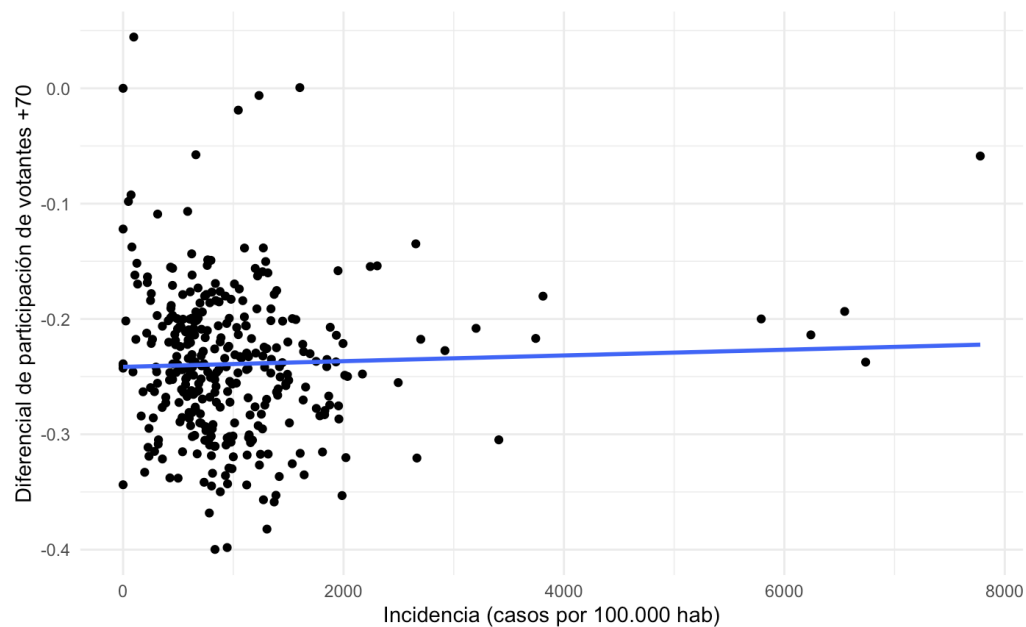
Fuente: Elaboración propia utilizando datos de distribución de votantes publicados por el SERVEL.

Figura 3.9: Relación entre diferencial de participación a nivel comunal y tasa de incidencia para votantes entre 50-59 años de edad en el plebiscito nacional de 2020



Fuente: Elaboración propia utilizando datos de distribución de votantes publicados por el SERVEL.

Figura 3.10: Relación entre diferencial de participación a nivel comunal y tasa de incidencia para votantes +70 años de edad en el plebiscito nacional de 2020



Fuente: Elaboración propia utilizando datos de distribución de votantes publicados por el SERVEL.

Tabla 3.5: Resumen de prueba Kruskal-Wallis para testear por diferencias significativas de cambios en participación entre grupos etarios

Test	Chi-squared	df	p-value
Kruskal-Wallis Rank Sum	1480	5	< 2.2e-16

Tabla 3.6: Resultados de pruebas t de Welch para comparar medias de cambios en participación en el plebiscito nacional entre comunas en base a *Fases12*

Age group	Mean 1	Mean 2	t	df	p-value
18-29	0.1238478	0.1770448	-5.7123	321	2.543e-08
30-39	0.02491598	0.07230547	-5.1536,	334.85	4.377e-07
40-49	-0.06692307	-0.02481032	-4.7737	337.06	2.7e-06
50-59	-0.1596316	-0.1112948	-4.656	332.52	4.667e-06
60-69	-0.2296359	-0.1926261	-3.8201	339.74	0.0001586
70+	-0.2433398	-0.2339114	-1.4107	324.91	0.1593

Tabla 3.7: Pruebas *Wilcoxon Rank Sum* para comparar cambios en participación en el plebiscito nacional entre comunas en base a *Fases12*

Age group	W-statistic	p-value
18-29	9540	4.908e-08
30-39	10030	8.808e-07
40-49	10318	4.219e-06
50-59	10310	4.044e-06
60-69	10736	3.453e-05
70+	12153	0.009727

Tabla 3.8: Resultados de pruebas t de Welch para comparar medias de cambios en participación en el plebiscito nacional entre comunas en base a *Incidencia*

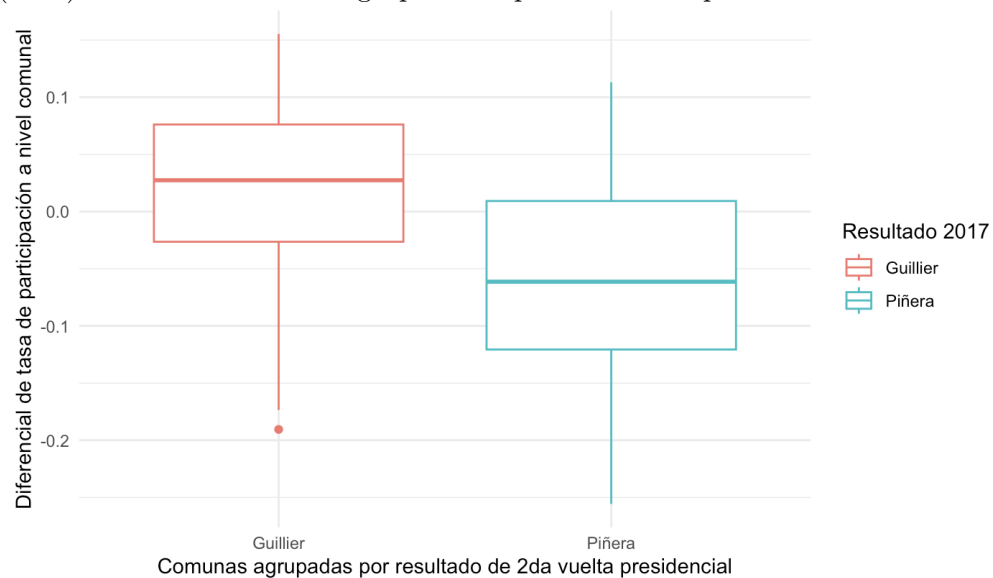
Age group	Mean 1	Mean 2	t	df	p-value
18-29	0.1425683	0.1536579	-1.1503	335.92	0.2508
30-39	0.03915376	0.05391071	-1.5504,	334.95	0.122
40-49	-0.05193941	-0.04348809	-0.92584	335.22	0.3552
50-59	-0.1385248	-0.1366417	-0.17616	0.17616	0.8603
60-69	-0.2105370	-0.2149714	0.4445	339.94	0.657
70+	-0.2321295	-0.2459488	2.0137	338.63	0.04483

Tabla 3.9: Pruebas *Wilcoxon Rank Sum* para comparar cambios en participación en el plebiscito nacional entre comunas en base a *Incidencia*

Age group	W-statistic	p-value
18-29	13637	0.2823
30-39	13213	0.1238
40-49	13834	0.3899
50-59	14492	0.8887
60-69	15150	0.5628
70+	16459	0.04438

Para analizar la misma variable de resultado considerando el “color-político” de cada comuna, se agrupan las zonas en base a los candidatos presidenciales que obtuvieron mayor votación en cada comuna durante la segunda vuelta presidencial de 2017. Se ve en la siguiente figura que la participación en 2020 incrementó en promedio en las comunas donde hubo más apoyo por Alejandro Guillier, mientras que las comunas con más apoyo por Sebastian Piñera participaron menos, posiblemente indicando una pérdida de votantes inclinados hacia la “derecha” en 2020.

Figura 3.11: Diferencial de participación entre el plebiscito de entrada (2020) y la elección presidencial anterior (2017) en comunas agrupadas por color político histórico



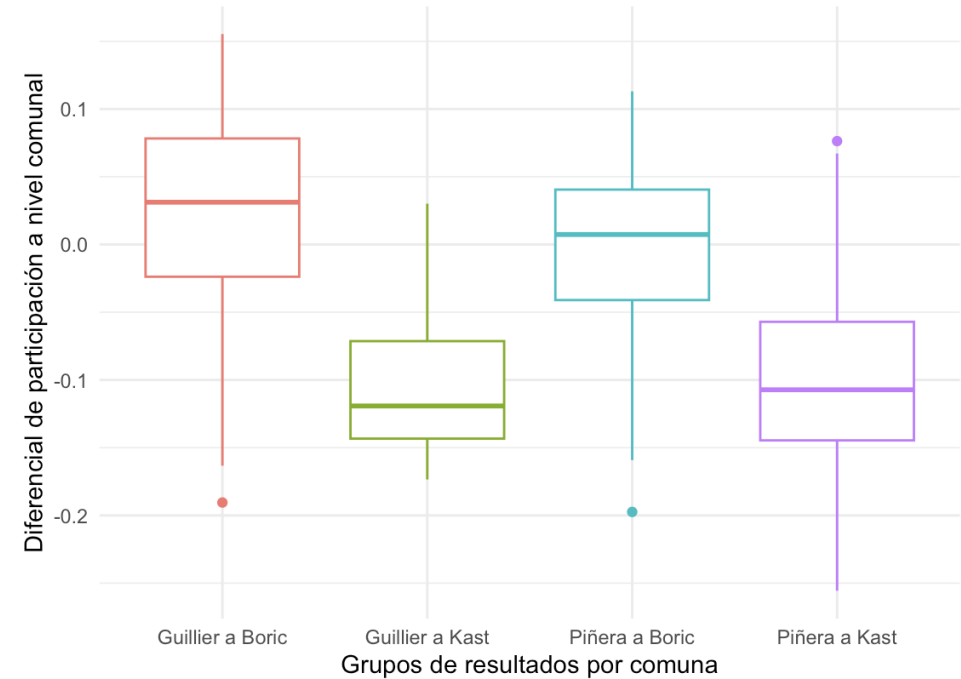
Fuente: Elaboración propia utilizando datos de participación publicados por el SERVEL.

En el gráfico que sigue, se desea mostrar el diferencial de participación en las comunas que cambiaron o mantuvieron su preferencia de tipo de candidato, donde un apoyo por Guillier y Kast en una misma comuna, junto a un apoyo por Piñera y Boric, se consideran como cambios de preferencia dada las diferencias de los candidatos en términos de posición ideológica. Se observa que hubo una disminución en participación en comunas que apoyaron primero a Guillier y luego Kast, posiblemente indicando que hubo un cambio de votantes en estas

comunas. Por otro lado, en las comunas donde hubo apoyo por ambos Piñera y Boric, se mantuvo en promedio la participación, señalando la posibilidad de que los mismos votantes pre-pandemia en estas comunas votaron de manera distinta durante la pandemia.

Por último, de acuerdo a la Tabla 3.10, la cual muestra la cantidad de mesas de votación donde cada par de candidatos obtuvo la mayor cantidad de votos, hubieron más de 10.000 mesas que pasaron de votar mayormente por Sebastián Piñera en 2017 a Gabriel Boric en la primera vuelta presidencial de 2021 lo cual se puede deber a una aparición de nuevos votantes o un cambio de preferencia por parte de votantes que previamente apoyaron a Sebastian Piñera.

Figura 3.12: Diferencial de participación de comunas agrupadas por candidatos presidenciales de 2017 y 2021



Fuente: Elaboración propia utilizando datos resultados y participación publicados por el SERVEL.

Tabla 3.10: Conteo de mesas ganadas por candidato utilizando resultados de elecciones presidenciales publicados por el Servel (2017 vs 2021)

2021	2017				
	GUILLIER	PIÑERA	SANCHEZ	KAST	GOIC
BORIC	3397	10125	5261	3	3
KAST	1216	17852	390	46	1
PARISI	426	1571	877	34	0
PROVOSTE	244	562	0	0	2
SICHEL	4	109	0	0	0
OMINAMI	12	12	4	0	0

3.4. Preparación de datos

La preparación de datos consiste en definir las variables de resultado y transformar las variables explicativas de acuerdo a los métodos a implementar, donde se desean especificar modelos con la unidad de análisis más granular posible, las mesas de votación. Primero, para poder identificar a las mesas entre los periodos de estudio y por ende construir los datos de panel, se crea un identificador para cada mesa. Dado que se fusionan las mesas de votación en cada elección en distintas combinaciones durante el proceso de conteo de votos, se realiza una homolagación de mesas. Con las mesas identificadas entre periodos esto permite también formatear las variables de estudio de manera *long* o *wide* según corresponde.

Luego, para la especificación de modelos asociados a **H1**, se debe considerar que no se disponen de la cantidad de inscritos por rango etario a nivel de mesa, por lo que no se pueden obtener las tasas de participación de cada grupo de votantes para la unidad de análisis. Por esta razón, se define la primera variable de resultado como la proporción de votantes entre M y N años en la mesa i para la elección t , $VotantesM_N_{it}$, normalizada por la cantidad de inscritos en el mismo rango en la comuna donde se ubica la mesa. Además, se le aplica una transformación *logit* o *log odds*, $\log\left(\frac{p}{1-p}\right)$, donde p es la proporción de votantes normalizada. Esta última transformación sirve para aproximar los datos a una distribución normal.

En cuanto a la definición de las variables de tratamiento, se establecen valores umbrales para cada uno de los factores COVID, $Incidencia_{it}$, $Fases12_{it}$ y $Fases123_{it}$, y se transforman a variables binarias para identificar a las mesas pertenecientes a comunas con alta incidencia y bajo un mayor tiempo en confinamiento. A diferencia de **H1**, los datos relacionados a **H2** corresponden a datos transversales que incluyen las mismas variables de tratamiento mencionadas anteriormente pues esta segunda hipótesis solo consiste en estudiar la primera vuelta presidencial de 2021.

3.5. Modelamiento

En la etapa de modelamiento se explican los métodos a aplicar para poder responder a las hipótesis planteadas. Para **H1**, se estiman modelos de diferencia-en-diferencias (*DiD*) para cada rango etario y cada variable de tratamiento descrita anteriormente donde se considera un intervalo de confianza del 95 % en la interpretación de resultados. En cuanto a **H2**, se ejecutan distintos modelos *logit* multinomiales para modelar las preferencias por mesa de votación durante la primera vuelta presidencial. La variable de resultado, $Choice_i$, es una variable categórica nominal que representa el candidato con mayor proporción de votos en cada mesa, donde las alternativas corresponden a los siguientes candidatos: *Boric*, *Kast*, *Parisi*, *Provoste*, *Sichel* y *Ominami*.

3.5.1. *DiD*

En las tablas 3.11 y 3.12 se muestran las estimaciones de las regresiones asociadas al plebiscito nacional de 2020 cuyas especificaciones corresponden a la ecuación 3.1. Se puede ver que el estimador *DiD* correspondiente al tratamiento de confinamiento ($Treated = 1$ si $Fases12 \geq 45$ días) es positivo y estadísticamente significativo para votantes entre 18-29 y 40-49, indicando que un mayor tiempo en confinamiento fue asociado a una mayor proporción de votantes en estos rangos etarios en comparación a mesas ubicadas en zonas con menor tiempo en confinamiento estricto. Particularmente, el *logit* de la proporción de votantes entre 18-29 años (normalizada) es 0.063 unidades mayor en las mesas tratadas, o equivalentemente, las *odds* de la proporción aumentan por un factor de $e^{0.063} \approx 1.07$, lo que representa un incremento del 7 % en las *odds*.

$$\log\left(\frac{Y_{it}}{1 - Y_{it}}\right) = \beta_0 + \beta_1 Treated_i + \beta_2 Post_t + \beta_3 (Treated_i \times Post_t) + \beta_4 Dens_i + \epsilon_{it} \quad (3.1)$$

En cuanto al tratamiento de incidencia, donde $Treated = 1$ si la mesa pertenece a una comuna con alta incidencia (*Incidencia* mayor a la mediana), el estimador *DiD* es significativo únicamente para el grupo de votantes entre 60-69 años, donde la intervención fue asociada a un efecto negativo en la proporción de votantes de este rango durante el plebiscito nacional en comparación a las mesas ubicadas en comunas con baja incidencia. Particularmente, se asoció a una disminución del $1 - e^{-0.044} * 100 \approx 4\%$ en las *odds* de la proporción.

Los modelos asociados a la segunda vuelta presidencial de 2021 son especificados de la misma manera con la diferencia de que el factor de confinamiento, *Fases123*, considera también el tiempo en la tercera fase del Plan Paso a Paso. Las estimaciones de estos modelos se pueden encontrar en los anexos A.5 y A.6. Las tablas indican que ninguno de los factores COVID incidieron en una mayor o menor proporción de votantes en las mesas tratadas de manera significativa durante la segunda vuelta presidencial, con excepción de uno de los grupos etarios, donde la alta incidencia fue asociado a una mayor proporción de votantes entre 40-49 ($DiD \approx 0.099$) en comparación a la proporción de votantes en mesas con baja incidencia.

Tabla 3.11: Resultados de estimación *DiD* para determinar significancia de *Fases12* en la proporción de votantes por grupo etario en el plebiscito nacional

Variable	Coefficient	Std. Error	t-value	P-value
18-29				
(Intercept)	-11.360407 ***	0.009943	-1142.592	<2e-16
Post	0.521458 ***	0.013920	37.460	<2e-16
Treatment (<i>Fases12</i>)	-1.336863 ***	0.012811	-104.353	<2e-16
Dens	-3.144e-05 ***	8.788e-07	-35.772	<2e-16
Post × Treatment (DiD)	0.063086 ***	0.017929	3.519	0.000434
Residual Std. Error	1.281 (Df = 85306)			
Adjusted <i>R</i> ²	0.2398			
30-39				
(Intercept)	-11.478407 ***	0.013342	-860.324	<2e-16
Post	0.286137 ***	0.018680	15.318	<2e-16
Treatment (<i>Fases12</i>)	-1.386553 ***	0.017191	-80.656	<2e-16
Dens	-3.636e-05 ***	1.181e-06	-30.770	<2e-16
Post × Treatment (DiD)	0.008202	0.024059	0.341	0.733
Residual Std. Error	1.719 (Df = 85306)			
Adjusted <i>R</i> ²	0.1488			
40-49				
(Intercept)	-10.93810 ***	0.01176	-930.088	<2e-16
Post	-0.29456 ***	0.01647	-17.890	<2e-16
Treatment (<i>Fases12</i>)	-1.38790 ***	0.01515	-91.592	<2e-16
Dens	-4.280e-05 ***	1.037e-06	-41.276	<2e-16
Post × Treatment (DiD)	0.06568 ***	0.02121	3.097	0.00195
Residual Std. Error	1.515 (Df = 85306)			
Adjusted <i>R</i> ²	0.182			
50-59				
(Intercept)	-10.61427 ***	0.01084	-978.867	<2e-16
Post	-0.17741 ***	0.01518	0.01518	<2e-16
Treatment (<i>Fases12</i>)	-1.37575 ***	0.01397	-98.467	<2e-16
Dens	-3.703e-05 ***	9.572e-07	-38.685	<2e-16
Post × Treatment (DiD)	0.03083	0.01955	1.577	0.115
Residual Std. Error	1.397 (Df = 85306)			
Adjusted <i>R</i> ²	0.2011			
60-69				
(Intercept)	-10.53574 ***	0.01169	-900.968	<2e-16
Post	-0.29036 ***	0.01637	-17.735	<2e-16
Treatment (<i>Fases12</i>)	-1.38070 ***	0.01507	-91.635	<2e-16
Dens	-3.532e-05 ***	1.034e-06	1.034e-06	<2e-16
Post × Treatment (DiD)	-0.02865	0.02109	-1.358	0.174
Residual Std. Error	1.507 (Df = 85306)			
Adjusted <i>R</i> ²	0.1886			
70+				
(Intercept)	-10.87902 ***	0.01279	-850.524	<2e-16
Post	-0.54492 ***	0.01791	-30.428	<2e-16
Treatment (<i>Fases12</i>)	-1.33613 ***	0.01648	-81.070	<2e-16
Dens	-3.078e-05 ***	1.134e-06	-27.144	<2e-16
Post × Treatment (DiD)	-0.02391	0.02307	-1.037	0.3
Residual Std. Error	1.648 (Df = 85306)			
Adjusted <i>R</i> ²	0.1663			
<i>Note:</i>		*p<0.1; **p<0.05; ***p<0.01		

Tabla 3.12: Resultados de estimación *DiD* para determinar significancia de *Incidencia* en la proporción de votantes por grupo etario en el plebiscito nacional

Variable	Coefficient	Std. Error	t-value	P-value
18-29				
(Intercept)	-11.86 ***	1.060e-02	-1118.679	<2e-16
Post	0.5377 ***	1.298e-02	41.425	<2e-16
Treatment (<i>Incidencia</i>)	-0.1199 ***	1.414e-02	-8.482	<2e-16
Dens	-7.230e-05 ***	9.733e-07	-74.277	<2e-16
Post × Treatment (DiD)	0.0355 *	1.900e-02	1.868	0.0618
Residual Std. Error	1.384 (Df = 85306)			
Adjusted R^2	0.09892			
30-39				
(Intercept)	-12.011 ***	1.382e-02	-869.017	<2e-16
Post	0.2785 ***	1.691e-02	16.467	<2e-16
Treatment (<i>Incidencia</i>)	-0.0822 ***	1.842e-02	-4.460	8.2e-06
Dens	-7.810e-05 ***	1.268e-06	-61.579	<2e-16
Post × Treatment (DiD)	0.01479	2.476e-02	0.597	0.55
Residual Std. Error	1.804 (Df = 85306)			
Adjusted R^2	0.05224			
40-49				
(Intercept)	-11.442 ***	1.224e-02	-934.463	<2e-16
Post	-0.2449 ***	1.499e-02	-16.340	<2e-16
Treatment (<i>Incidencia</i>)	-0.09742 ***	1.632e-02	-5.969	2.4e-09
Dens	-8.439e-05 ***	1.124e-06	-75.099	<2e-16
Post × Treatment (DiD)	-0.03441	2.194e-02	-1.568	0.117
Residual Std. Error	1.598 (Df = 85306)			
Adjusted R^2	0.07175			
50-59				
(Intercept)	-11.180 ***	1.146e-02	-975.318	<2e-16
Post	-0.1467 ***	1.403e-02	-10.456	<2e-16
Treatment (<i>Incidencia</i>)	-0.01241	1.528e-02	-0.813	0.4164
Dens	-7.646e-05 ***	1.052e-06	-72.700	<2e-16
Post × Treatment (DiD)	-0.03801 *	2.054e-02	-1.851	0.0641
Residual Std. Error	1.496 (Df = 85306)			
Adjusted R^2	0.06811			
60-69				
(Intercept)	-11.130 ***	1.231e-02	-904.085	<2e-16
Post	-0.02928 ***	1.506e-02	-19.437	<2e-16
Treatment (<i>Incidencia</i>)	0.02958 *	1.641e-02	1.803	0.0714
Dens	-7.464e-05 ***	1.130e-06	-66.079	<2e-16
Post × Treatment (DiD)	-0.04365 **	2.205e-02	-1.979	0.0478
Residual Std. Error	1.606 (Df = 85306)			
Adjusted R^2	0.06488			
70+				
(Intercept)	-11.471 ***	1.331e-02	-861.514	<2e-16
Post	-0.05477 ***	1.629e-02	-33.626	<2e-16
Treatment (<i>Incidencia</i>)	0.03457 *	1.774e-02	1.949	0.0513
Dens	-6.883e-05 ***	1.221e-06	-56.349	<2e-16
Post × Treatment (DiD)	-0.03592	2.385e-02	-1.506	0.1321
Residual Std. Error	1.737 (Df = 85306)			
Adjusted R^2	0.06558			
<i>Note:</i>		*p<0.1; **p<0.05; ***p<0.01		

3.5.2. *Logit multinomial*

Próximamente se estiman y evalúan modelos logit multinomiales para analizar el comportamiento de elección de los votantes durante la primera vuelta presidencial. El primer modelo especificado, cuyas estimaciones se resumen en la Tabla 3.13, sirve para determinar si efectivamente existió un cambio en preferencia, particularmente por parte de votantes que apoyaron al candidato Sebastian Piñera en la primera vuelta presidencial de 2017. Para distinguir entre mesas de votación con o sin nuevos votantes, se crea un identificador llamado *new voters*, el cual indica si una mesa tuvo un incremento en participación con respecto a la elección anterior o si mantuvo una participación similar junto a un cambio en composición etaria (*new voters* = 0 para mesas que no cumplen con ninguno de estos dos casos). Además, se crea el indicador *pinera* para identificar a las mesas donde ganó Sebastián Piñera en 2017. Cada columna de coeficientes en la tabla se asocia a cada alternativa/candidato de la elección de 2021. Cabe destacar que no aparece una columna para la opción de *boric* dado que esta es configurada como la opción base, es decir, todas las estimaciones de coeficientes representan el efecto en elegir a cada alternativa en lugar de elegir a *boric*.

El efecto particular de interés se define en la expresión 3.2, el efecto de *new voters* en las mesas que previamente votaron *pinera* en elegir a *kast* por sobre *boric*. El valor de este efecto corresponde a la suma de los coeficientes del término de interacción y de *new voters* asociados a *kast*, $0.177 + -0.520 = -0.343$. El signo de este valor indica que: para el universo de mesas que previamente apoyaron mayormente a *pinera*, las mesas con nuevos votantes fueron menos probables de elegir *kast* sobre *boric* que las mesas sin nuevos votantes. Siendo coeficientes significativos, la interpretación anterior indica que no hubo una preferencia por Boric por parte de los votantes que previamente apoyaron a *pinera* en 2017.

$$E(\pi(\frac{y = kast}{y = boric}) \mid new\ voters = 1, pinera = 1) - \quad (3.2)$$

$$E(\pi(\frac{y = kast}{y = boric}) \mid new\ voters = 0, pinera = 1)$$

Si bien no existió un cambio en preferencia general, se puede analizar si hubo un cambio heterogéneo en base a factores COVID. En la Tabla 3.14, se resume otra especificación del modelo la cual incluye *high incid* en lugar de *new voters*, una variable que indica si una mesa pertenece a una comuna con alta incidencia de COVID-19. De acuerdo a las estimaciones de este modelo, dentro del universo de mesas que previamente apoyaron a *pinera*, las mesas con alta incidencia fueron menos probables de votar por *kast* sobre *boric* que mesas con baja incidencia. A pesar de que este resultado indica un cambio de preferencia heterogéneo por parte de las mesas que votaron *pinera*, es posible que el efecto sea mayormente influido por las mesas con nuevos votantes.

Para estimar la expresión 3.3 la cual diferencia el efecto de alta incidencia por mesas con nuevos votantes y votantes similares, se especifica un tercer modelo que incluye la variable de *new voters* y cuyos resultados se ven en la Tabla 3.15. Al sumar los coeficientes de todas las interacciones asociadas a *new voters*, junto con su efecto individual ($-0.865 + 0.386 + 0.908 - 0.515 = -0.086$), se obtiene que para el universo de mesas que previamente apoyaron a *pinera*, aquellas mesas con alta incidencia y con nuevos votantes fueron menos probables

de votar *kast* sobre *boric* que mesas con alta incidencia y votantes similares.

$$E(\pi(\frac{y = kast}{y = boric}) \mid new\ voters = 1, pinera = 1, high\ incid = 1) - \quad (3.3)$$

$$E(\pi(\frac{y = kast}{y = boric}) \mid new\ voters = 0, pinera = 1, high\ incid = 1)$$

Tabla 3.13: Resumen de logit multinomial de la primera vuelta presidencial de 2021 (ML1)

	<i>Dependent variable:</i>				
	kast	ominami	parisi	provoste	sichel
	(1)	(2)	(3)	(4)	(5)
pinera	2.061*** (0.069)	0.235 (1.237)	-0.186 (0.126)	0.459*** (0.133)	2.827*** (1.023)
new voters	-0.520*** (0.069)	0.887 (1.043)	0.390*** (0.100)	-1.288*** (0.137)	-0.805 (1.158)
pinera:new voters	0.177** (0.076)	-0.800 (1.301)	0.235* (0.133)	0.205 (0.165)	0.385 (1.180)
Constant	-1.220*** (0.063)	-7.044*** (1.013)	-2.215*** (0.095)	-2.565*** (0.112)	-7.027*** (1.004)
Akaike Inf. Crit.	75,792.450	75,792.450	75,792.450	75,792.450	75,792.450

Note:

*p<0.1; **p<0.05; ***p<0.01

A continuación se realiza el mismo análisis anterior para el tratamiento de confinamiento *fases123*, el cual indica si una mesa pertenece al grupo de comunas que estuvieron una mayor cantidad de días bajo las restricciones de las primeras tres fases del Plan Paso a Paso. Como se puede ver en la Tabla 3.16, la suma de los coeficientes de *fases123* y del término de interacción ($0.094 + -0.437 = -0.343$) indica que las mesas de votación que estuvieron bajo mayor tiempo en confinamiento y que votaron *pinera* en 2017 fueron menos probables de elegir *kast* sobre *boric* que aquellas mesas con menor tiempo en restricciones. Además, diferenciando por mesas con votantes nuevos y votantes similares (Tabla 3.17), las mesas con nuevos votantes y mayor tiempo en confinamiento fueron menos probables en elegir *kast* sobre *boric* que las mesas con votantes similares, donde el efecto total de *new voters* es igual a $-0.230 - 0.099 - 0.465 + 0.511 = -0.283$.

Tabla 3.14: Resumen de logit multinomial de la primera vuelta presidencial de 2021 (ML2)

	<i>Dependent variable:</i>				
	kast	ominami	parisi	provoste	sichel
	(1)	(2)	(3)	(4)	(5)
pinera	2.248*** (0.041)	-0.588 (0.517)	0.181*** (0.052)	0.678*** (0.096)	2.322*** (0.604)
high incid	-1.490*** (0.058)	-1.093** (0.493)	-1.411*** (0.062)	-1.613*** (0.142)	-1.831 (1.154)
pinera:high incid	0.051 (0.064)	0.242 (0.760)	-0.435*** (0.087)	0.111 (0.169)	1.814 (1.173)
Constant	-0.918*** (0.035)	-5.636*** (0.317)	-1.152*** (0.039)	-2.786*** (0.078)	-6.842*** (0.578)
Akaike Inf. Crit.	71,561.050	71,561.050	71,561.050	71,561.050	71,561.050

Note:

*p<0.1; **p<0.05; ***p<0.01

Tabla 3.16: Resumen de logit multinomial de la primera vuelta presidencial de 2021 (ML4)

	<i>Dependent variable:</i>				
	kast	ominami	parisi	provoste	sichel
	(1)	(2)	(3)	(4)	(5)
pinera	2.452*** (0.052)	-8.719*** (0.191)	-0.287*** (0.056)	0.474*** (0.091)	2.961*** (0.587)
fases123	0.094 (0.059)	1.983* (1.031)	-0.810*** (0.059)	-1.653*** (0.141)	-1.883 (1.154)
pinera:fases123	-0.437*** (0.064)	8.556*** (0.191)	0.408*** (0.081)	-0.514*** (0.196)	-1.231 (1.241)
Constant	-1.722*** (0.049)	-7.903*** (1.001)	-1.387*** (0.043)	-2.749*** (0.078)	-6.806*** (0.578)
Akaike Inf. Crit.	75,307.430	75,307.430	75,307.430	75,307.430	75,307.430

Note:

*p<0.1; **p<0.05; ***p<0.01

Tabla 3.15: Resumen de logit multinomial de la primera vuelta presidencial de 2021 (ML3)

	<i>Dependent variable:</i>				
	kast	ominami	parisi	provoste	sichel
	(1)	(2)	(3)	(4)	(5)
pinera	1.902*** (0.097)	-4.721 (8.187)	-0.170 (0.163)	0.562*** (0.168)	2.361** (1.033)
new voters	-0.865*** (0.093)	0.122 (1.059)	0.242* (0.132)	-1.397*** (0.172)	-1.375 (1.228)
high incid	-2.250*** (0.150)	-6.889 (19.533)	-1.562*** (0.201)	-1.718*** (0.242)	-5.837 (11.575)
pinera:new voters	0.386*** (0.107)	4.314 (8.204)	0.415** (0.172)	-0.041 (0.207)	-0.354 (1.277)
pinera:high incid	0.505*** (0.163)	10.904 (21.167)	-0.151 (0.269)	-0.526* (0.300)	4.274 (11.583)
new voters:high incid	0.908*** (0.163)	5.934 (19.539)	0.174 (0.211)	0.074 (0.300)	4.436 (11.640)
pinera:new voters:high incid	-0.515*** (0.177)	-11.198 (21.183)	-0.322 (0.284)	1.052*** (0.366)	-2.224 (11.651)
Constant	-0.195** (0.084)	-5.746*** (1.005)	-1.370*** (0.126)	-1.679*** (0.143)	-5.747*** (1.005)
Akaike Inf. Crit.	71,044.370	71,044.370	71,044.370	71,044.370	71,044.370

Note:

*p<0.1; **p<0.05; ***p<0.01

Tabla 3.17: Resumen de logit multinomial de la primera vuelta presidencial de 2021 (ML5)

	<i>Dependent variable:</i>				
	kast	ominami	parisi	provoste	sichel
	(1)	(2)	(3)	(4)	(5)
pinera	2.518*** (0.115)	5.007*** (0.612)	-0.361** (0.170)	0.390** (0.156)	2.464** (1.020)
new voters	-0.230* (0.122)	11.987*** (0.710)	0.704*** (0.140)	-0.953*** (0.166)	-0.844 (1.224)
fases123	0.502*** (0.133)	13.234*** (0.710)	-0.420** (0.191)	-1.238*** (0.248)	-19.377*** (0.730)
pinera:new voters	-0.099 (0.129)	-21.525*** (0.645)	0.103 (0.180)	0.076 (0.193)	0.662 (1.246)
pinera:fases123	-0.860*** (0.145)	-4.521*** (0.612)	0.317 (0.253)	-0.555* (0.330)	17.109*** (0.730)
new voters:fases123	-0.465*** (0.149)	-11.401*** (0.733)	-0.461** (0.201)	-0.452 (0.303)	17.810*** (0.730)
pinera:new voters:fases123	0.511*** (0.162)	20.808*** (0.645)	0.094 (0.267)	-0.020 (0.412)	-18.966*** (0.840)
Constant	-1.537*** (0.109)	-19.692*** (0.710)	-1.997*** (0.132)	-2.045*** (0.135)	-6.168*** (0.999)
Akaike Inf. Crit.	74,873.200	74,873.200	74,873.200	74,873.200	74,873.200

Note:

*p<0.1; **p<0.05; ***p<0.01

3.6. Evaluación

Llegando a la última etapa del proceso *CRISP-DM*, se evalúan los resultados de los modelos estimados anteriormente utilizando las métricas de evaluación correspondientes. La métrica utilizada para evaluar las regresiones lineales es la R^2 -ajustado la cual indica el porcentaje de la variabilidad en la variable de resultado que es explicada por el modelo considerando el número de predictores, donde un valor más cercano a 1 es mejor. La evaluación de los modelos *DiD* es que resultaron en valores de R^2 -ajustado bajos, indicando bajo poder explicativo y estimadores posiblemente sesgados en cuanto a sus magnitudes. Sin embargo, la significancia de los estimadores permiten concluir sobre la direccionalidad de los efectos.

En general, el factor de confinamiento tuvo mejor poder explicativo para el plebiscito nacional que para la segunda vuelta presidencial. Particularmente, tuvo efectos significativos positivos en la proporción de votantes entre 18-29 y 40-49 en el plebiscito, explicando el 24 % y el 18 % de la variabilidad en las proporciones de votantes para estos rangos de edad respectivamente. Por otro lado, el factor de incidencia tuvo mejor poder explicativo para la segunda vuelta presidencial que para el plebiscito nacional, donde la alta incidencia tuvo un efecto positivo significativo en la proporción de votantes entre 40-49, explicando el 12 % de la variabilidad de la proporción.

Tabla 3.18: Valores de R-cuadrado ajustado para modelos DiD asociados al plebiscito nacional por tratamiento

Modelo/Grupo	con <i>Incidencia</i>	con <i>Fases12</i>
18-29	0.0989	0.2398
30-39	0.0522	0.1488
40-49	0.0718	0.182
50-59	0.0681	0.2011
60-69	0.0649	0.187
70+	0.0656	0.166

Tabla 3.19: Valores de R-cuadrado ajustado para modelos DiD asociados a segunda vuelta presidencial por tratamiento

Modelo/Grupo	con <i>Incidencia</i>	con <i>Fases123</i>
18-29	0.148	0.115
30-39	0.0892	0.0719
40-49	0.123	0.09764
50-59	0.141	0.107
60-69	0.128	0.0960
70+	0.101	0.0739

Cabe destacar que para poder inferir causalidad con respecto a los efectos de los factores, los grupos de tratamiento y control deben cumplir con el supuesto de tendencias paralelas, lo cual implica que, en ausencia del tratamiento (COVID-19), la diferencia entre los grupos habría permanecido constante a lo largo del tiempo. Para poder verificar el cumplimiento del supuesto de manera estadística, se estiman regresiones que permiten determinar si existieron diferencias significativas entre las mesas tratadas y no-tratadas en el periodo pre-pandemia (2017). Los resultados de las regresiones, los cuales se pueden ver en los anexos A.7 a A.10, revelan que sí existieron diferencias significativas entre los grupos para ambos tratamientos y todos los rangos etarios. Por lo tanto, no se puede inferir efectos causales de los factores COVID estudiados.

Las métricas de evaluación de los modelos *logit* multinomiales incluyen el criterio AIC (*Akaike Information Criterion*) el cual evalúa la calidad de ajuste de cada modelo versus su complejidad. El valor absoluto de la métrica no tiene una interpretación directa a diferencia del R^2 -ajustado utilizado para regresiones con variable de resultado continua. Por lo tanto, para realizar una evaluación de los modelos de elección discreta estimados previamente bajo la misma muestra de mesas de votación, se deben comparar los valores AIC de cada modelo, donde un valor menor indica un mejor modelo. De acuerdo a la Tabla 3.20, el modelo mejor evaluado relativo al modelo *logit* 1 es el modelo *logit* 3, pues al incorporar el término de interacción entre las variables *pinera*, *high incid* y *new voters*, el AIC baja de 75,792 a 71,044. Además, los coeficientes de interés en los demás modelos poseen errores estándares altos, indicando baja precisión de las estimaciones.

Tabla 3.20: Valores *AIC* para la evaluación de modelos *logit* multinomiales ML1 a ML5

	ML 1	ML 2	ML 3	ML 4	ML 5
Valor AIC	75,792	71,561	71,044	75,307	74,873

Capítulo 4

CONCLUSIONES

A partir de los resultados obtenidos anteriormente, se concluye primero que las mesas de votación ubicadas en zonas con alta incidencia de COVID-19 se asociaron con una menor proporción de votantes entre 60-69 años durante el plebiscito nacional de 2020 y una mayor porcentaje de votantes entre 40-49 años en la segunda vuelta presidencial de 2021 en comparación a aquellas mesas ubicadas en comunas con baja incidencia, implicando que no hay suficiente evidencia para respaldar lo planteado en la hipótesis **H1b** para los demás grupos de votantes mayores en ambas elecciones. Además, zonas bajo un mayor tiempo en las fases más estrictas del Plan Paso a Paso tuvieron efectos positivos en la proporción de votantes entre 18-29 y 40-49 años en el plebiscito nacional, confirmando solo parcialmente lo planteado en la hipótesis **H1a**. Estos resultados fueron obtenidos a partir del método de diferencia-en-diferencias considerando un intervalo de confianza del 95 %, sin embargo, los modelos sólo explican entre el 5 % al 24 % de la variabilidad de las variables de resultado estudiadas: la proporción de votantes en cada rango etario a nivel de mesa. Además, como ningún modelo cumple con el supuesto de tendencias paralelas, no se puede inferir efectos causales de los factores de incidencia y confinamiento.

En seguida se concluye sobre los modelos de elección discreta estimados para determinar si hubieron cambios de preferencia de candidato asociados a los factores COVID-19 durante la primera vuelta presidencial en pandemia, específicamente, si ciudadanos ubicados en zonas más impactadas por el virus y que previamente apoyaron a Sebastian Piñera en 2017, un candidato de “derecha”, votaron por Gabriel Boric, un candidato de “izquierda”, en 2021. Las estimaciones de los *logit* multinomiales indican que aquellas mesas con votantes similares a la elección anterior y que previamente apoyaron mayormente a Sebastián Piñera fueron más probables de votar por José Kast sobre Gabriel Boric en comparación a las mesas con votantes nuevos. Se obtuvo el mismo resultado incorporando las variables de tratamiento de incidencia de COVID-19 y tiempo en confinamiento, indicando que estos factores no incidieron en un cambio de preferencia durante la primera vuelta presidencial y que el apoyo por Gabriel Boric en mesas con una tendencia política histórica de derecha fue influido mayormente por las preferencias de votantes nuevos, refutando lo planteado en la hipótesis **H2**.

Para futuros trabajos, se propone primero otra especificación de la variable de resultado estudiada en el caso de poder obtener los datos necesarios a nivel de mesa de votación: nivel de participación por rango etario, en lugar de proporción de votantes por rango. Segundo, se propone estudiar otra elección que también ocurrió en pandemia con una tasa de partici-

pación nacional del 42 % - la primera elección de convencionales constituyentes - analizando si los factores COVID incidieron en las preferencias por partido político. Por último, sería interesante analizar el efecto de otras variables relacionadas a COVID-19 más complejas de obtener como por ejemplo la tasa de contagio-mortalidad o la tasa de hospitalizaciones a nivel de comuna.

Bibliografía

- Adam-Troian, J., Bonetto, E., Varet, F., Arciszewski, T., and Guiller, T. (2023). Pathogen threat increases electoral success for conservative parties: Results from a natural experiment with covid-19 in france. *Evolutionary Behavioral Sciences*, 17(3):357.
- Giommoni, T. and Loumeau, G. (2020). Lockdown and voting behaviour: A natural experiment on postponed elections during the covid-19 pandemic. *SSRN Electronic Journal*.
- Karwowski, M., Kowal, M., Groyecka, A., Białek, M. J., Lebuda, I., Sorokowska, A., and Sorokowski, P. (2020). When in danger, turn right: Does covid-19 threat promote social conservatism and right-wing presidential candidates? *Human Ethology*.
- Mondschein, S. V., Olivares, M. A., Ordóñez, F., Schwartz, D., Weintraub, A., Torres-Ulloa, I., Aguayo, C., and Canessa, G. (2022). Service design to balance waiting time and infection risk: An application for elections during the covid-19 pandemic. *Service Science*, 14(2):90–109.
- Scheller, D. S. (2021). Pandemic primary: the interactive effects of covid-19 prevalence and age on voter turnout. *Journal of Elections, Public Opinion and Parties*, 31:180 – 190.
- Warshaw, C., Vavreck, L., and Baxter-King, R. (2020). Fatalities from covid-19 are reducing americans’ support for republicans at every level of federal office. *Science Advances*, 6.

ANEXOS

Anexo A

Anexo A.1. Estadísticas descriptivas segunda vuelta presidencial 2017

Tabla 4.1: Estadísticas descriptivas: características *mesas y comunas* correspondientes a la **segunda vuelta presidencial de 2017**

Statistic	N	Mean	St. Dev.	Min	Max
<i>Participacion</i>	41,736	0.49	0.11	0.01	0.97
<i>Mujer</i>	41,736	0.53	0.27	0.00	1.00
<i>Hombre</i>	41,736	0.47	0.27	0.00	1.00
<i>Votantes18_29</i>	41,736	0.19	0.24	0.0000	1.00
<i>Votantes30_39</i>	41,736	0.15	0.15	0.0004	0.97
<i>Votantes40_49</i>	41,736	0.17	0.14	0.0004	0.98
<i>Votantes50_59</i>	41,736	0.20	0.10	0.0005	0.90
<i>Votantes60_69</i>	41,736	0.16	0.09	0.0003	0.75
<i>Votantes+70</i>	41,736	0.12	0.09	0.0003	0.73
<i>PromInscMesa</i>	41,736	326.96	27.79	30.10	348.00
<i>Inscritos18_29</i>	346	0.03	0.01	0.01	0.06
<i>Inscritos30_39</i>	346	0.18	0.03	0.11	0.31
<i>Inscritos40_49</i>	346	0.19	0.02	0.10	0.34
<i>Inscritos50_59</i>	346	0.18	0.01	0.09	0.23
<i>Inscritos60_69</i>	346	0.12	0.01	0.06	0.17
<i>Inscritos+70</i>	346	0.12	0.04	0.02	0.54

Anexo A.2. Matriz de correlación de variables asociadas al plebiscito nacional

Tabla 4.2: Matriz de correlación de variables COVID-19 para el plebiscito nacional de 2020

	Fases12	Incidencia	Densidad
Fases12	1	0.114	0.331
Incidencia	0.114	1	-0.265
Densidad	0.331	-0.265	1

Anexo A.3. Matriz de correlación de variables asociadas a la segunda vuelta presidencial

Tabla 4.3: Matriz de correlación de variables COVID-19 para la segunda vuelta presidencial de 2021

	Fases123	Incidencia	Densidad
Fases123	1	0.590	0.337
Incidencia	0.590	1	0.100
Densidad	0.337	0.100	1

Anexo A.4. Pruebas Levene para testear por homogeneidad de varianza de datos asociados al plebiscito nacional a nivel de comuna

Tabla 4.4: Resumen de pruebas Levene para testear por homogeneidad de varianza

Age group	Test Statistic (F)	Df	p-value
18-29	1.8512	1, 340	0.1745
30-39	3.3016	1, 340	0.07009
40-49	3.509	1, 340	0.0619
50-59	0.038	1, 340	0.8455
60-69	0.1123	1, 340	0.7377
70+	0.485	1, 340	0.4866

Anexo A.5. Resumen de modelos DiD asociados a la segunda vuelta presidencial con el factor confinamiento

Tabla 4.5: Resultados de estimación *DiD* para determinar significancia de *Fases123* en la proporción de votantes por grupo etario en la segunda vuelta presidencial

Variable	Coefficient	Std. Error	t-value	P-value
18-29				
(Intercept)	-11.550 ***	1.240e-02	-931.745	<2e-16
Post	0.4028 ***	1.726e-02	23.338	<2e-16
Treatment (<i>Fases123</i>)	-0.6185 ***	1.511e-02	-40.940	<2e-16
Dens	-4.834e-05 ***	9.252e-07	-52.250	<2e-16
Post × Treatment (DiD)	0.03924 *	2.026e-02	1.937	0.0528
Residual Std. Error		1.336 (Df = 87455)		
Adjusted R^2		0.1153		
30-39				
(Intercept)	-11.661 ***	1.631e-02	-715.091	<2e-16
Post	0.2452 ***	2.271e-02	10.799	<2e-16
Treatment (<i>Fases123</i>)	-0.6453 ***	1.988e-02	-32.464	<2e-16
Dens	-5.508e-05 ***	1.217e-06	-45.249	<2e-16
Post × Treatment (DiD)	0.04401 *	2.666e-02	1.651	0.0988
Residual Std. Error		1.758 (Df = 87455)		
Adjusted R^2		0.07197		
40-49				
(Intercept)	-11.090 ***	1.392e-02	1.392e-02	<2e-16
Post	-0.2073 ***	1.938e-02	-10.701	<2e-16
Treatment (<i>Fases123</i>)	-0.6540 ***	1.696e-02	-38.560	<2e-16
Dens	-6.049e-05 ***	1.039e-06	-58.230	<2e-16
Post × Treatment (DiD)	0.08673 ***	2.275e-02	3.812	0.000138
Residual Std. Error		1.5 (Df = 87455)		
Adjusted R^2		0.09764		
50-59				
(Intercept)	-10.770 ***	1.301e-02	-827.991	<2e-16
Post	-0.1303 ***	1.811e-02	-7.196	6.27e-13
Treatment (<i>Fases123</i>)	-0.6744 ***	1.585e-02	-42.546	<2e-16
Dens	-5.514e-05 ***	9.708e-07	-56.804	<2e-16
Post × Treatment (DiD)	0.03276	2.126e-02	1.541	0.123
Residual Std. Error		1.402 (Df = 87455)		
Adjusted R^2		0.1068		
60-69				
(Intercept)	-0.1068 ***	1.398e-02	-763.982	<2e-16
Post	-0.1742 ***	1.947e-02	-8.946	<2e-16
Treatment (<i>Fases123</i>)	-0.6954 ***	1.704e-02	-40.805	<2e-16
Dens	-5.400e-05 ***	1.044e-06	-51.740	<2e-16
Post × Treatment (DiD)	0.02923	2.286e-02	1.279	0.201
Residual Std. Error		1.508 (Df = 87455)		
Adjusted R^2		0.09603		
70+				
(Intercept)	-11.070 ***	1.543e-02	-717.346	<2e-16
Post	-0.2261 ***	2.148e-02	-10.528	<2e-16
Treatment (<i>Fases123</i>)	-0.6149 ***	0.01880	-32.706	<2e-16
Dens	-5.227e-05 ***	1.151e-06	-45.396	<2e-16
Post × Treatment (DiD)	-9.882e-03	2.522e-02	-0.392	0.695
Residual Std. Error		1.663 (Df = 87455)		
Adjusted R^2		0.07393		
<i>Note:</i>				
*p<0.1; **p<0.05; ***p<0.01				

Anexo A.6. Resumen de modelos DiD asociados a la segunda vuelta presidencial con el factor de incidencia

Tabla 4.6: Resultados de estimación *DiD* para determinar significancia de *Incidencia* en la proporción de votantes por grupo etario en la segunda vuelta presidencial

Variable	Coefficient	Std. Error	t-value	P-value
18-29				
(Intercept)	-11.350 ***	1.280e-02	-886.567	<2e-16
Post	0.4066 ***	1.782e-02	22.822	<2e-16
Treatment (<i>Incidencia</i>)	-0.8850 ***	1.512e-02	-58.526	<2e-16
Dens	-4.526e-05 ***	8.868e-07	-51.033	<2e-16
Post \times Treatment (DiD)	0.03405 *	2.055e-02	1.657	0.0975
Residual Std. Error		1.312 (Df = 87455)		
Adjusted R^2		0.1478		
30-39				
(Intercept)	-11.500 ***	1.700e-02	-676.360	<2e-16
Post	0.2775 ***	2.366e-02	11.726	<2e-16
Treatment (<i>Incidencia</i>)	-0.8579 ***	2.008e-02	-42.718	<2e-16
Dens	-5.301e-05 ***	1.178e-06	-45.012	<2e-16
Post \times Treatment (DiD)	7.620e-04	2.729e-02	0.028	0.978
Residual Std. Error		1.742 (Df = 87455)		
Adjusted R^2		0.08919		
40-49				
(Intercept)	-10.890 ***	1.443e-02	-754.417	<2e-16
Post	-0.2181 ***	2.009e-02	-10.854	<2e-16
Treatment (<i>Incidencia</i>)	-0.9193 ***	1.705e-02	-53.912	<2e-16
Dens	-5.773e-05 ***	1.000e-06	-57.726	<2e-16
Post \times Treatment (DiD)	9.920e-02 ***	2.317e-02	4.281	1.86e-05
Residual Std. Error		1.479 (Df = 87455)		
Adjusted R^2		0.1231		
50-59				
(Intercept)	-10.550 ***	1.342e-02	-786.414	<2e-16
Post	-0.1286 ***	1.868e-02	-6.883	5.9e-12
Treatment (<i>Incidencia</i>)	-0.9548 ***	1.586e-02	-60.216	<2e-16
Dens	-5.211e-05 ***	9.299e-07	-56.042	<2e-16
Post \times Treatment (DiD)	3.069e-02	2.154e-02	1.425	0.154
Residual Std. Error		1.375 (Df = 87455)		
Adjusted R^2		0.1407		
60-69				
(Intercept)	-10.460 ***	1.445e-02	-724.187	<2e-16
Post	-0.1642 ***	2.012e-02	-8.162	3.33e-16
Treatment (<i>Incidencia</i>)	-0.9772 ***	1.707e-02	-57.230	<2e-16
Dens	-5.093e-05 ***	1.001e-06	-50.863	<2e-16
Post \times Treatment (DiD)	0.01635	2.320e-02	0.705	0.481
Residual Std. Error		1.481 (Df = 87455)		
Adjusted R^2		0.1275		
70+				
(Intercept)	-10.830 ***	1.598e-02	-677.779	<2e-16
Post	-0.2139 ***	2.225e-02	-9.610	<2e-16
Treatment (<i>Incidencia</i>)	-0.9238 ***	1.889e-02	-48.910	<2e-16
Dens	-4.814e-05 ***	1.108e-06	-43.454	<2e-16
Post \times Treatment (DiD)	-0.02441	2.567e-02	-0.951	0.342
Residual Std. Error		1.638 (Df = 87455)		
Adjusted R^2		0.1014		
<i>Note:</i> *p<0.1; **p<0.05; ***p<0.01				

Anexo A.7. Resumen de pruebas de tendencias paralelas para modelos asociados al plebiscito nacional con el factor de confinamiento

Tabla 4.7: Prueba de tendencias paralelas pre-tratamiento (*Fases12*)

Variable	Coefficient	Std. Error	t-value	P-value
18-29				
(Intercept)	-11.36041 ***	0.01088	-1043.68	<2e-16
Treatment (<i>Fases12</i>)	-1.33686 ***	0.01403	-95.32	<2e-16
Residual Std. Error	1.402 (Df = 41734)			
Adjusted R^2	0.1788			
40-49				
(Intercept)	-10.93810 ***	0.01101	-993.48	<2e-16
Treatment (<i>Fases12</i>)	-1.38790 ***	0.01419	-97.83	<2e-16
Residual Std. Error	1.418 (Df = 41734)			
Adjusted R^2	0.1865			
50-59				
(Intercept)	-10.61427 ***	0.01095	-969.41	<2e-16
Treatment (<i>Fases12</i>)	-1.37575 ***	0.01411	-97.52	<2e-16
Residual Std. Error	1.411 (Df = 41734)			
Adjusted R^2	0.1856			
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01			

Anexo A.8. Resumen de pruebas de tendencias paralelas para modelos asociados al plebiscito nacional con el factor de incidencia

Tabla 4.8: Prueba de tendencias paralelas pre-tratamiento (*Incidencia* correspondiente a 2020)

Variable	Coefficient	Std. Error	t-value	P-value
18-29				
(Intercept)	-12.24551 ***	0.01035	-1182.9	<2e-16
Treatment (<i>Incidencia</i>)	0.17131 ***	0.01516	11.3	<2e-16
Residual Std. Error	1.545 (Df = 41734)			
Adjusted R^2	0.003026			
50-59				
(Intercept)	-11.58074 ***	0.01043	-1110.82	<2e-16
Treatment (<i>Incidencia</i>)	0.29559 ***	0.01527	19.36	<2e-16
Residual Std. Error	1.556 (Df = 41734)			
Adjusted R^2	0.008876			
60-69				
(Intercept)	-11.52135 ***	0.01093	-1054.18	<2e-16
Treatment (<i>Incidencia</i>)	0.33024 ***	0.01601	20.63	<2e-16
Residual Std. Error	1.631 (Df = 41734)			
Adjusted R^2	0.01007			
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01			

Anexo A.9. Resumen de pruebas de tendencias paralelas para modelos asociados a la segunda vuelta presidencial con el factor de confinamiento

Tabla 4.9: Prueba de tendencias paralelas pre-tratamiento (*Fases123*)

Variable	Coefficient	Std. Error	t-value	P-value
18-29				
(Intercept)	-11.56268 ***	0.01395	-829.04	<2e-16
Treatment (<i>Fases123</i>)	-0.83014 ***	0.01638	-50.69	<2e-16
Residual Std. Error	1.504 (Df = 42320)			
Adjusted R^2	0.05722			
30-39				
(Intercept)	-11.67745 ***	0.01774	-658.12	<2e-16
Treatment (<i>Fases123</i>)	-0.88645 ***	0.02083	-42.55	<2e-16
Residual Std. Error	1.913 (Df = 42320)			
Adjusted R^2	0.041			
40-49				
(Intercept)	-11.10770 ***	0.01414	-785.80	<2e-16
Treatment (<i>Fases123</i>)	-0.91882 ***	0.01660	-55.36	<2e-16
Residual Std. Error	1.524 (Df = 42320)			
Adjusted R^2	0.06751			
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01			

Anexo A.10. Resumen de pruebas de tendencias paralelas para modelos asociados a la segunda vuelta presidencial con el factor de incidencia

Tabla 4.10: Prueba de tendencias paralelas pre-tratamiento (*Incidencia* correspondiente a 2021)

Variable	Coefficient	Std. Error	t-value	P-value
18-29				
(Intercept)	-11.37233 ***	0.01443	-788.21	<2e-16
Treatment (<i>Incidencia</i>)	-1.05475 ***	0.01664	-63.37	<2e-16
Residual Std. Error	1.48 (Df = 42320)			
Adjusted R^2	0.08664			
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01			