



Universidad del Desarrollo
Facultad de Ingeniería

ESTRATIFICACIÓN DE LOS PERMISOS DE MOVILIDAD EN SANTIAGO DE
CHILE DURANTE LA PANDEMIA EN 2020

POR: CAMILO RIVERA VERDUGO

Proyecto de grado presentado a la Facultad de Ingeniería de la Universidad del
Desarrollo para optar al grado académico de Magíster en Data Science

PROFESOR GUÍA:

Sr. LEONARDO FERRES

DICIEMBRE 2025

SANTIAGO

TABLA DE CONTENIDO

RESUMEN	1
1. INTRODUCCIÓN	2
2. TRABAJO RELACIONADO	4
3. HIPÓTESIS Y OBJETIVOS	6
4. DATOS Y METODOLOGÍA	7
4.1. DATOS	7
4.1.1. <i>Estructura del panel comuna-día</i>	9
4.2. METODOLOGÍA.....	10
5. RESULTADOS	13
5.1. ANÁLISIS EXPLORATORIO DE DATOS (EDA)	13
5.2. MODELOS DE REGRESIÓN LINEAL CON EFECTOS FIJOS	26
5.3. RANDOM FOREST COMO MODELO NO LINEAL DE REFERENCIA	30
5.4. XGBOOST COMO MODELO DE GRADIENTE BOOSTING.....	33
6. CONCLUSIONES	36
BIBLIOGRAFÍA	38

Resumen

La pandemia de Covid-19 en Chile introdujo restricciones sanitarias inéditas, incluyendo cuarentenas territoriales y un sistema de permisos individuales para desplazarse durante los periodos de confinamiento. Estas políticas generaron un “experimento” sobre la movilidad de las personas, que puede analizarse combinando datos administrativos y fuentes masivas de movilidad derivadas de la red celular.

En este trabajo se estudia en qué medida la movilidad observada a nivel de comuna, medida a partir de registros de telefonía móvil (XDR), puede explicarse por la movilidad autorizada a través de permisos temporales. Para ello se construye un panel comuna–día para la Región Metropolitana durante 2020, combinando: (i) los índices de movilidad interna y externa del dataset publicado por Pappalardo et al. (2023); (ii) los registros diarios de permisos emitidos por comuna; (iii) la información de períodos de cuarentena por comuna; y (iv) datos censales de población para calcular tasas per cápita.

Sobre este panel se realiza un análisis exploratorio temporal y espacial de la movilidad y los permisos por 1.000 habitantes, y se estiman modelos de regresión lineal con controles de calendario y efectos fijos por comuna, complementados con modelos de aprendizaje supervisado (Random Forest y XGBoost) entrenados sobre la misma estructura comuna–día.

En conjunto, los modelos sugieren que la mayor parte de la variación en los índices de movilidad se explica por diferencias estructurales entre comunas y por el calendario de la pandemia, mientras que el sistema de permisos captura solo parcialmente los cambios reales en la movilidad. Estos resultados aportan evidencia cuantitativa sobre las limitaciones de los permisos como instrumento de control de movilidad.

1. Introducción

La pandemia de Covid-19 modificó de manera importante los patrones de movilidad y uso del espacio público en Chile. A partir de marzo de 2020 se implementó un conjunto de restricciones sanitarias que incluyeron cuarentenas territoriales, toques de queda y un sistema de permisos individuales para desplazarse, gestionado en línea. Estas medidas alteraron la vida cotidiana de millones de personas y abrieron preguntas sobre su grado de cumplimiento y sus efectos reales en la movilidad.

Paralelamente, en los últimos años se han desarrollado fuentes de datos masivos que permiten observar el movimiento de las personas a partir de señales de la red celular. En particular, Pappalardo et al. (2023) publican un conjunto de índices de movilidad derivados de registros XDR (eXtended Detail Records) para Chile durante 2020, desagregados por comuna y día. Estos índices permiten medir la intensidad de la movilidad interna dentro de la comuna y la movilidad externa hacia otras comunas, ofreciendo información sobre cómo cambiaron los patrones de desplazamiento durante la primera ola de la pandemia.

Este trabajo busca conectar ambos mundos: La movilidad autorizada por el sistema de permisos y la movilidad observada vía XDR, enfocándose en la Región Metropolitana de Santiago. La idea central es evaluar en qué medida los permisos emitidos por comuna y día logran explicar los cambios en los índices de movilidad, una vez que se controla por factores de calendario y por la heterogeneidad estructural entre comunas.

El resto del documento se organiza de la siguiente manera. La Sección 2 revisa la literatura relacionada sobre la medición de movilidad a partir de datos de telefonía móvil, el impacto de las restricciones sanitarias y el uso de datos administrativos en la evaluación de políticas públicas. La Sección 3 presenta la pregunta de investigación, la hipótesis y los objetivos específicos del estudio. La Sección 4 describe en detalle las fuentes de datos, la construcción del panel comuna–día y la metodología de análisis, incluyendo modelos de

panel con efectos fijos y modelos de aprendizaje supervisado. La Sección 5 presenta los resultados, comenzando por el análisis exploratorio de la relación entre movilidad, permisos y cuarentenas, y continuando con las estimaciones de regresión y los modelos de Random Forest y XGBoost. Finalmente, la Sección 6 resume las conclusiones y sugiere líneas de trabajo futuro.

2. Trabajo Relacionado

En la última década, el uso de datos de telefonía móvil —registros de detalle de llamadas (CDR) y registros extendidos (XDR)— se ha consolidado como una herramienta relevante para medir la movilidad humana y apoyar decisiones de salud pública. Diversas revisiones muestran que estos datos permiten construir indicadores agregados de desplazamiento, monitorear patrones de viaje, evaluar el impacto de intervenciones no farmacéuticas (NPIs) y apoyar la modelación epidemiológica en tiempo casi real, aunque enfrentan desafíos de representatividad, gobernanza y transparencia metodológica (Oliver et al., 2020).

Para el caso chileno, Pappalardo et al. (2023) publican un conjunto de índices de movilidad derivados de XDR para el año 2020, desagregados por comuna y día, que incluyen movilidad interna (IM_int), externa (IM_ext) y total (IM). Este dataset se diseñó explícitamente para evaluar el impacto de las cuarentenas locales en Chile y se ha transformado en un insumo de referencia para estudios sobre la efectividad de las restricciones sanitarias y la heterogeneidad territorial en la respuesta de movilidad frente a las NPIs.

En paralelo, existe literatura sobre el diseño institucional de las restricciones y el rol de los permisos de desplazamiento. En Chile, Araya-Moreno (2022) analiza los permisos como artefactos burocráticos con efectos materiales (control, fiscalización), pero típicamente sin una medición granular y comparable de la movilidad efectivamente realizada en todas las comunas, lo que limita cuantificar el alineamiento entre movilidad “autorizada” y movilidad observada.

El presente trabajo se ubica en esa brecha y propone combinar el dataset público de índices de movilidad basados en XDR con registros administrativos de permisos de desplazamiento y períodos de cuarentena por comuna, con el fin de cuantificar la relación entre permisos, movilidad real y heterogeneidades territoriales en la Región

Metropolitana. Para identificar esta relación explotando la variación intra-comuna en el tiempo, se especifica un modelo de datos de panel con efectos fijos por comuna (y controles temporales) que permite absorber heterogeneidad no observada constante en el tiempo —por ejemplo, atributos territoriales, estructura urbana o patrones persistentes de actividad— siguiendo el tratamiento estándar de la econometría de panel (Wooldridge, 2010; Baltagi, 2005).

3. Hipótesis y Objetivos

3.1 Hipótesis

La implementación del pase de movilidad en el contexto de las cuarentenas de fases durante la pandemia de COVID-19 tiene un impacto significativo en la movilización de la población y el éxito o fracaso de las estrategias de control de la pandemia.

3.2 Objetivos

Objetivo general: Cuantificar la asociación entre los permisos de movilidad por 1.000 habitantes y los índices de movilidad intercomunal observados a partir de datos XDR en la Región Metropolitana durante 2020, considerando el calendario de la pandemia y el régimen de cuarentenas por comuna.

Objetivos específicos

1. Construir un panel comuna–día para la Región Metropolitana combinando índices de movilidad XDR, permisos de desplazamiento, información de cuarentenas y datos censales de población, generando variables per cápita y de calendario.
2. Describir la evolución temporal y la heterogeneidad territorial de la movilidad y de los permisos por 1.000 habitantes, identificando patrones por comuna, día de la semana, mes y períodos con cuarentena
3. Estimar modelos de regresión lineal (OLS) con efectos fijos por comuna y controles de calendario para evaluar la asociación entre la tasa de permisos y la movilidad externa observada, evaluando el aporte marginal de los permisos frente a las cuarentenas y a la estructura comunal.
4. Entrenar modelos Random Forest y XGBoost para modelar para modelar la movilidad externa a partir de permisos, cuarentenas y controles de calendario, con el fin de capturar posibles relaciones no lineales e interacciones no representadas explícitamente en la regresión lineal.

5. Comparar el desempeño predictivo de la regresión lineal, Random Forest y XGBoost utilizando métricas como R2, RMSE y MAE y analizar la importancia de las variables en los modelos de árbol, poniendo énfasis en rol de los permisos, las cuarentenas y los efectos territoriales.

4. Datos y Metodología

4.1. Datos

El análisis se basa en la construcción de un panel comuna–día para la Región Metropolitana de Santiago durante 2020. Para ello se integran las siguientes fuentes de datos:

1. **Índices de movilidad (XDR).** Dataset `mobility_indices.csv`, derivado del trabajo de Pappalardo et al. (2023), que contiene un panel comuna–día para todo Chile. Para cada comuna (`cid`) y día (`date`) entre el 26-02-2020 y el 20-09-2020, el archivo incluye tres índices de movilidad: interna, externa y total (`IM_interno`, `IM_externo`, `IM`). Estos índices se construyeron a partir de registros de detalle extendido (XDR) de la red celular y miden, de forma agregada, el número promedio de desplazamientos diarios por usuario de telefonía móvil residente en la comuna, distinguiendo viajes dentro de la misma comuna y hacia otras comunas. El dataset incorpora además identificadores territoriales (`region`, `rid`, `comuna`, `cid`) y la superficie de la comuna (`area`). Para este estudio se trabaja únicamente con la Región Metropolitana, filtrando las observaciones con identificador de región `rid = 13`.
2. **Permisos de movilidad.** Registros de permisos emitidos por fecha, tipo de trámite, región y comuna. Estos datos se agregan a nivel de comuna–día para obtener el número total de permisos (`permisos_total`) y, combinados con la población

comunal censal, se construye la tasa de permisos por 1.000 habitantes (permisos_1000hab).

3. **Cuarentenas locales.** Archivo `quarantines.csv`, derivado del trabajo de Pappalardo et al. (2023), que registra para cada comuna la fecha y hora de inicio y término de cada período de cuarentena, junto con una descripción de su cobertura territorial. A partir de estos intervalos construimos un panel comuna–día y un indicador binario `en_cuarentena`, que toma valor 1 si la comuna estuvo en cuarentena en algún momento de ese día calendario. En esta construcción se consideran los horarios de inicio y levantamiento (frecuentemente nocturnos o de madrugada), de modo que los días de comienzo y término se clasifican coherentemente como días en cuarentena o no.
4. **Población censal 2024.** Archivo obtenido del Instituto Nacional de Estadísticas (INE), con población censada por comuna en 2024. Se filtra la Región Metropolitana y se utiliza la población como denominador para construir indicadores per cápita.
5. **Mapas vectoriales de comunas de Chile.** Archivo obtenido de la Biblioteca del Congreso Nacional, para incorporar análisis espacial en el análisis exploratorio de los datos.

A partir de estas fuentes se genera un panel rectangular de 52 comunas y 208 días, cubriendo el período desde el 26 de febrero de 2020 hasta el 20 de septiembre de 2020 (10.816 observaciones). Cada fila del panel corresponde a una combinación única comuna–día, lo que permite aplicar métodos de análisis de datos de panel.

En el proceso de construcción del panel se homologan los nombres de comuna entre las distintas fuentes, se filtran las observaciones con información inconsistente o incompleta y se verifica la ausencia de duplicados a nivel comuna–día. Esto asegura que el análisis exploratorio y los modelos posteriores se basen en un conjunto de datos coherente y comparables en todas las dimensiones.

4.1.1. Estructura del panel comuna-día

La Tabla 1 resume la estructura del panel final y describe las variables que lo componen. Todas ellas están alineadas a nivel de comuna-día y provienen de las fuentes descritas anteriormente o de transformaciones derivadas.

Tabla 1: Estructura del panel y descripción de variables.

Variable	Tipo	Descripción	Fuente principal
region	Catégorica (texto)	Nombre de la región (“Metropolitana de Santiago” en todas las filas).	mobility_indices.csv
rid	Entero	Código numérico de región (13 para la Región Metropolitana).	mobility_indices.csv
comuna	Catégorica (texto)	Nombre de la comuna, armonizado entre las distintas fuentes (por ejemplo, “Santiago”).	mobility_indices.csv + homologación con permisos y censo
cid	Entero	Código de comuna utilizado en el dataset de movilidad (identificador único por comuna).	mobility_indices.csv
area	Real	Superficie de la comuna (por ejemplo, en km ²).	mobility_indices.csv
IM_interno	Real	Índice de movilidad interna: intensidad de desplazamientos cuyo origen y destino están en la misma comuna, promediados por usuario.	mobility_indices.csv
IM_externo	Real	Índice de movilidad externa: intensidad de desplazamientos entre comunas, promediados por usuario.	mobility_indices.csv
IM	Real	Índice de movilidad total; en la práctica cumple aproximadamente que $IM \approx IM_interno + IM_externo$.	mobility_indices.csv
date	Fecha	Fecha de la observación (entre 2020-02-26 y 2020-09-20), a nivel de día calendario.	mobility_indices.csv / cálculo propio
permisos_total	Real	Número total de permisos emitidos en la comuna y día correspondientes (sumando todos los tipos de trámite).	Agregación diaria de Permisos emitidos
permisos_1000hab	Real	Tasa de permisos diarios por 1.000 habitantes, calculada como $permisos_total / (Población\ censada / 1000)$.	Permisos emitidos + Censo 2024
en_cuarentena	Entero (0/1)	Indicador de cuarentena: 1 si la comuna se encuentra bajo alguna forma de cuarentena en ese día, 0 si no.	quarantines.csv (intervalos expandidos a diario)
dow	Entero	Día de la semana (0 = lunes, ..., 6 = domingo), derivado de la fecha.	Cálculo propio a partir de date
es_finde	Entero (0/1)	Indicador de fin de semana: 1 si el día es sábado o domingo, 0 en caso contrario.	Cálculo propio a partir de dow
mes	Entero	Mes del año (2–9 en el período analizado), derivado de la fecha.	Cálculo propio a partir de date
anio	Entero	Año de la observación (2020 para todas las filas del panel).	Cálculo propio a partir de date
Población censada	Entero	Población censada de la comuna; es constante en el tiempo dentro de cada cid.	Censo 2024 (INE)

Fuente: Elaboración propia.

Esta estructura permite realizar análisis descriptivos temporales y espaciales de la movilidad y los permisos, y por otra parte aplicar modelos de regresión y aprendizaje

supervisado que explotan tanto la variación entre comunas como la variación dentro de cada comuna a lo largo del tiempo.

4.2. Metodología

La metodología del trabajo se organiza en tres etapas principales:

4.2.1. Análisis exploratorio de datos (EDA).

En una primera etapa se realiza un análisis descriptivo de los datos combinados en el panel comuna–día:

- Se describieron las distribuciones de los índices de movilidad interna, externa y total (IM, IM_externo, IM_interno) y de las variables de permisos (permisos_total, permisos_1000hab), tanto a nivel global como por comuna.
- Se construyeron series de tiempo agregadas para la Región Metropolitana y para comunas seleccionadas, comparando la evolución de la movilidad y de los permisos a lo largo del tiempo, y relacionándolas con hitos del calendario de la pandemia.
- Se analizaron diferencias entre días con y sin cuarentena y heterogeneidades espaciales entre comunas mediante gráficos de barras, mapas y correlaciones simples entre movilidad, permisos y fracción de días en cuarentena.

Esta etapa permite caracterizar la magnitud y variabilidad de la movilidad y de los permisos, e identificar posibles patrones temporales o territoriales para utilizar como información base en los modelos posteriores.

4.2.2. Modelos de regresión lineal con efectos fijos.

En la segunda etapa se estimaron modelos de regresión lineal (OLS) sobre el panel comuna-día, utilizando como variable dependiente el índice de movilidad externa $IM_{externo_{it}}$.

Se implementaron tres modelos progresivamente más robustos para abordar la heterogeneidad y los desafíos de inferencia estadística inherentes a los datos de panel:

4.2.2.1. Modelo OLS Agrupado (Pooled)

Inicialmente, se estimó un modelo OLS agrupado sin efectos fijos por comuna. Este modelo trata cada observación como independiente y sirve como punto de referencia ingenuo. Incluye controles de calendario mediante variables categóricas para el día de la semana y el mes:

$$IM_{externo_{it}} = \alpha + \beta \cdot permisos_{1000hab_{it}} + \gamma \cdot en_{cuarentena_{it}} + \delta_{dow} + \varphi_{mes} + \varepsilon_{it}$$

Sin embargo, este enfoque puede generar sesgo por omisión de variables si existen diferencias estructurales constantes entre las comunas (ej. densidad poblacional, accesibilidad geográfica, nivel socioeconómico, etc.) que influyen en la movilidad y en la implementación de políticas.

4.2.2.2. Modelos con Efectos Fijos por Comuna

Para mitigar el sesgo de heterogeneidad no observada y capturar cualquier característica constante en el tiempo que sea única para cada comuna, se incorporaron efectos fijos por comuna (α_i). De esta forma, la estimación del impacto de los permisos y las cuarentenas se basa exclusivamente en la variación observada dentro de cada comuna a lo largo del tiempo, neutralizando las diferencias promedio que hay entre ellas.

- **Modelo reducido:** Se enfoca únicamente en el efecto de la cuarentena, los efectos de calendario y los fijos de cada comuna:

$$IM_{externo_{it}} = \alpha_i + \gamma \cdot en_{cuarentena_{it}} + \delta_{dow} + \varphi_{mes} + \varepsilon_{it}$$

- **Modelo completo:** Agrega la tasa de permisos diarios, permitiendo evaluar su contribución adicional al modelo reducido:

$$IM_externo_{it} = \alpha_i + \beta \cdot \mathbf{permisos_1000hab}_{it} + \gamma \cdot en_cuarentena_{it} + \delta_{dow} + \varphi_{mes} + \varepsilon_{it}$$

En todos los casos, los coeficientes se estiman por MCO, pero la inferencia se realiza utilizando errores estándar robustos agrupados por comuna (`cov_type = "cluster"`, agrupando por `cid`), para permitir heterocedasticidad y correlación de los errores dentro de cada comuna en el tiempo. La contribución explicativa de `permisos_1000hab` se evalúa comparando el modelo reducido y el modelo completo a través de cambios en el Coeficiente de Determinación (R^2) asociado a la tasa de permisos.

4.2.3. Modelos de aprendizaje supervisado: Random Forest y XGBoost.

En la tercera etapa se utilizan modelos de aprendizaje supervisado como complemento de la regresión lineal, aprovechando su capacidad para capturar relaciones no lineales e interacciones complejas entre variables.

- Se formulan modelos de regresión supervisada donde la variable objetivo es nuevamente `IM_externo`, y las variables explicativas incluyen `permisos_1000hab`, `en_cuarentena`, los efectos de calendario (`dow`, `mes`) y el identificador de comuna (`cid`).
- Se emplea un esquema de **partición temporal** de los datos, utilizando los primeros meses del período de estudio como conjunto de entrenamiento y los meses posteriores como conjunto de prueba, de modo de respetar el orden temporal de la serie y evitar fugas de información.
- Sobre este esquema se ajustan un Random Forest Regressor y un XGBoost Regressor. Los modelos se configuran inicialmente con hiperparámetros estándar de la literatura y, en una etapa posterior, se realizan ajustes básicos (por ejemplo, número de árboles, profundidad máxima y tasa de aprendizaje) para mejorar el desempeño y la interpretación de los resultados.

El desempeño de los modelos se evalúa con métricas de error (RMSE, MAE) y coeficiente de determinación (R^2) en el conjunto de prueba. En el caso de los modelos de árbol, se analiza además la importancia de las variables para identificar qué factores explican en mayor medida la movilidad externa desde el punto de vista predictivo, poniendo énfasis en el rol de los permisos, por sobre las demás variables. Finalmente, se comparan los resultados de los modelos de regresión lineal y de aprendizaje supervisado, analizando las ventajas y limitaciones de cada enfoque en el contexto del problema de estudio.

5. Resultados

5.1. Análisis exploratorio de datos (EDA)

Antes de estimar modelos, se realiza un análisis exploratorio de la relación entre movilidad observada, permisos de desplazamiento y cuarentenas locales en la Región Metropolitana (RM) durante 2020. En primer lugar, se examinan las distribuciones de los índices de movilidad y de los permisos per cápita en el panel comuna–día.

A nivel agregado de RM, se observa el quiebre asociado al inicio de la pandemia:

- En febrero de 2020, la movilidad externa promedio (IM_externo) se sitúa en torno a 6,2 unidades, sin permisos registrados.
- El primer día con permisos distintos de cero es el 22 de marzo, mientras que las primeras cuarentenas comunales en la RM se observan desde el 26 de marzo.
- A partir de abril la movilidad externa promedio cae por debajo de 4,5, alcanzando su mínimo en mayo–junio ($\approx 4,0$), justo cuando aumenta la proporción de comunas en cuarentena y la intensidad de permisos.

En términos mensuales, la Tabla 2 resume la evolución de la movilidad externa promedio, la tasa de permisos por 1.000 habitantes y la fracción de comunas en cuarentena:

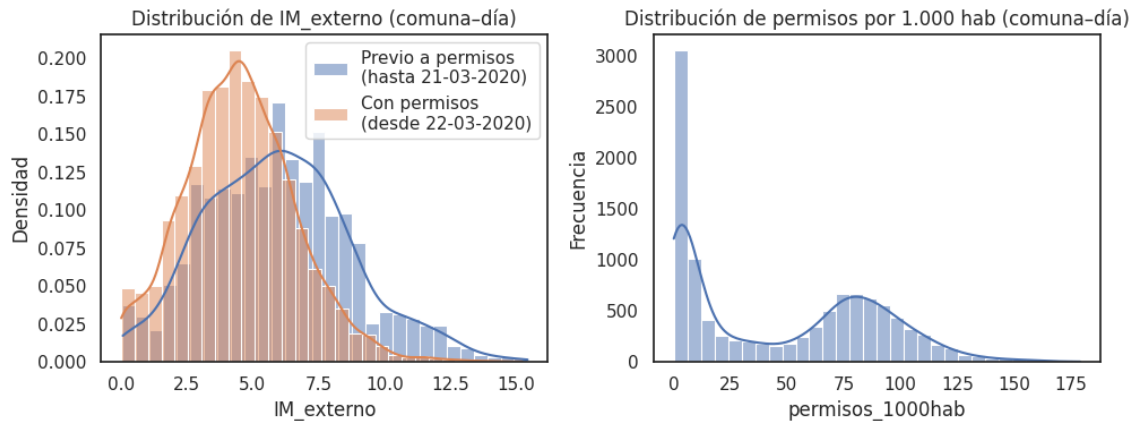
Tabla 2. Evolución mensual de movilidad, permisos y cuarentenas (RM, 2020)
(valores medios sobre comunas y días)

Mes	IM_externo_mean	permisos_1000hab_mean	en_cuarentena_share
2	6.18	0.00	0.00
3	5.29	3.43	0.03
4	4.21	14.64	0.13
5	4.04	46.89	0.56
6	4.00	71.96	0.80
7	4.51	78.05	0.87
8	5.08	75.93	0.68
9	5.82	45.65	0.32

Fuente: Elaboración propia.

La siguiente figura muestra las distribuciones de la movilidad externa (IM_externo) y de la tasa de permisos por 1.000 habitantes (permisos_1000hab) en el panel comuna-día de la Región Metropolitana. En el lado izquierdo se sobreponen las distribuciones de IM_externo para el periodo previo a la introducción de permisos (hasta el 21 de marzo de 2020) y para el periodo con permisos (desde el 22 de marzo en adelante).

Figura 1: Distribuciones de movilidad externa (IM_externo) y permisos por 1.000 habitantes a nivel comuna-día, Región Metropolitana (2020)



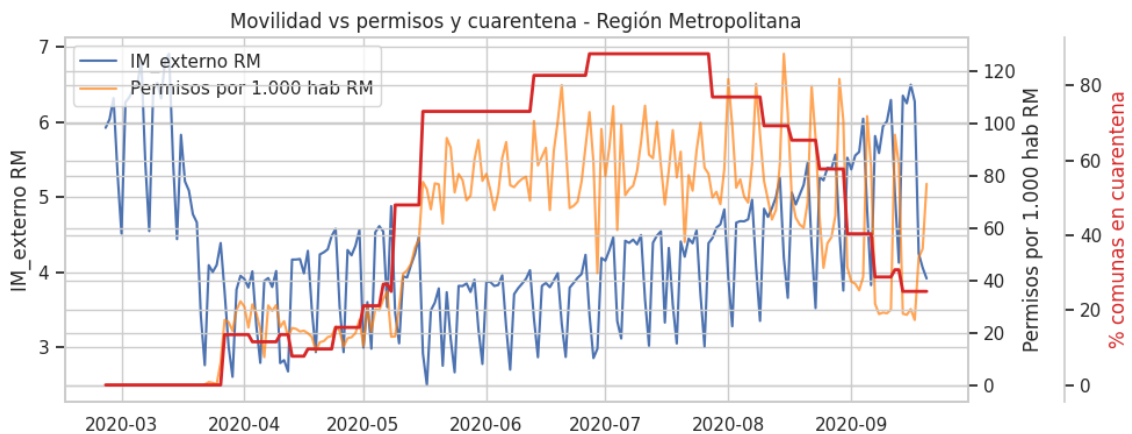
Fuente: Elaboración propia

Antes de los permisos, la movilidad externa presenta una distribución desplazada hacia valores más altos, con una media cercana a 6,0 unidades y un rango intercuartílico aproximado entre 4,0 y 7,7. En cambio, durante el periodo con permisos la distribución se desplaza hacia la izquierda, con una media alrededor de 4,5 y un rango intercuartílico entre 3,1 y 5,9, lo que refleja una caída importante pero no extrema de la movilidad externa tras la implementación de las medidas sanitarias. El panel derecho muestra que la distribución de permisos por 1.000 habitantes es altamente asimétrica: cerca de un 12% de las observaciones corresponde a días sin permisos (valor 0), mientras que el resto se concentra en un amplio rango que va desde valores bajos en los primeros meses hasta máximos cercanos a 180 permisos por 1.000 habitantes en los meses de mayor confinamiento.

Movilidad y permisos a nivel agregado regional

Para visualizar la dinámica conjunta de movilidad y permisos, se construye una serie de tiempo con el promedio diario en la RM de IM_externo y permisos_1000hab, junto con la proporción de comunas en cuarentena.

Figura 2: Movilidad externa y permisos por 1.000 habitantes en la Región Metropolitana (serie diaria, 2020).



Fuente: Elaboración propia.

En la Figura 2 (RM agregada) se aprecia que la movilidad externa exhibe un patrón claramente descendente a partir de mediados de marzo, coincidiendo con la declaración de la pandemia y las primeras cuarentenas. Durante las semanas de mayor confinamiento (mayo–julio) la movilidad se mantiene claramente por debajo del nivel prepandemia, sin retornar al valor de referencia de febrero dentro de la ventana observada.

En cambio, la tasa promedio de permisos por 1.000 habitantes muestra un rápido crecimiento, pasando de valores prácticamente nulos en marzo a alrededor de 70–80 permisos por 1.000 hab. en junio–julio, para luego disminuir gradualmente hacia septiembre. Los periodos donde más del 50% de las comunas están en cuarentena se asocian a los peaks de emisión de permisos y a los mínimos locales de movilidad externa,

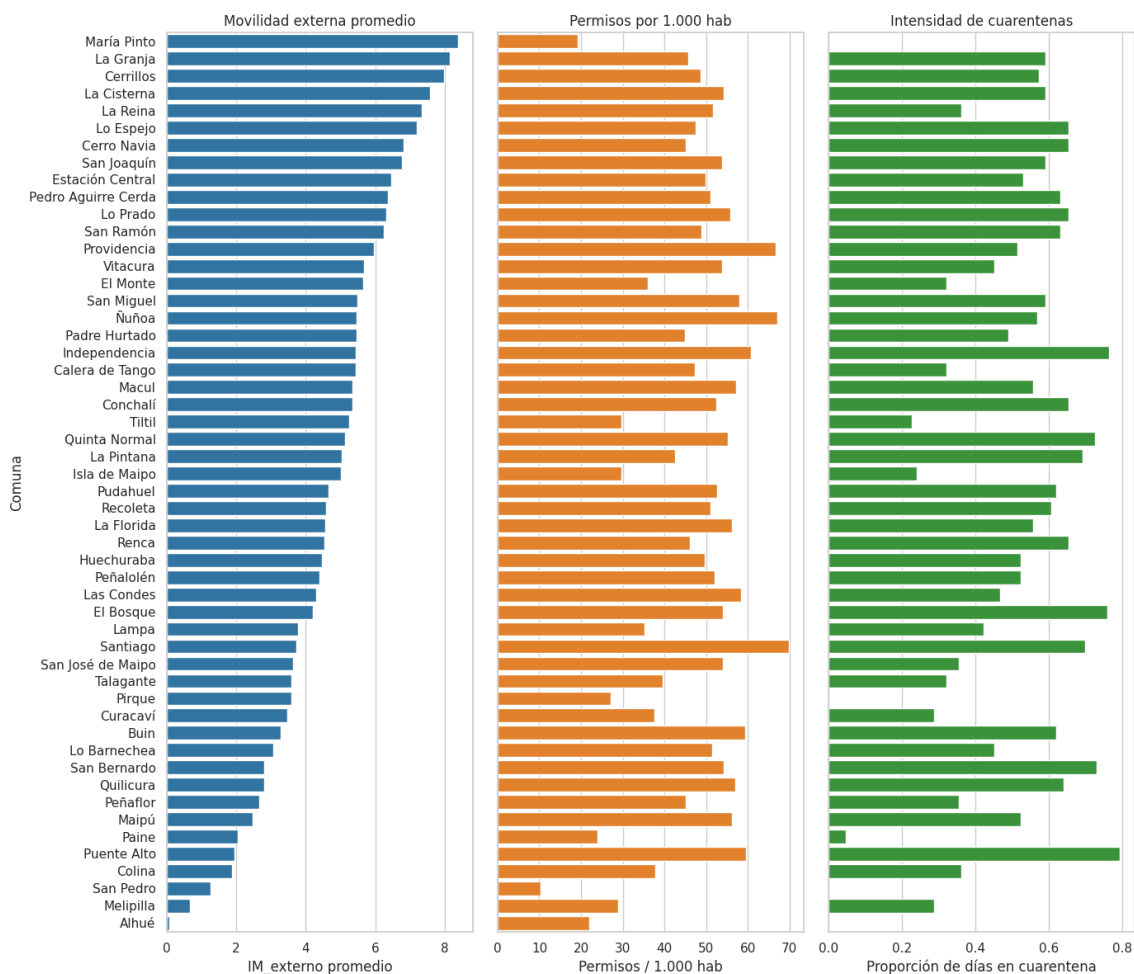
aunque la caída de la movilidad es mucho menos pronunciada que el incremento en permisos.

En resumen, el sistema de permisos se activa intensamente durante los meses de mayor confinamiento, mientras que la movilidad externa promedio se reduce, pero no en la misma magnitud que el aumento en permisos, lo que ya sugiere una relación no trivial entre ambas variables.

Heterogeneidad comunal

A continuación se exploran las diferencias entre comunas, tanto en niveles de movilidad y permisos como en intensidad de cuarentenas:

Figura 3: Heterogeneidad comunal de movilidad externa, permisos y cuarentenas en la Región Metropolitana

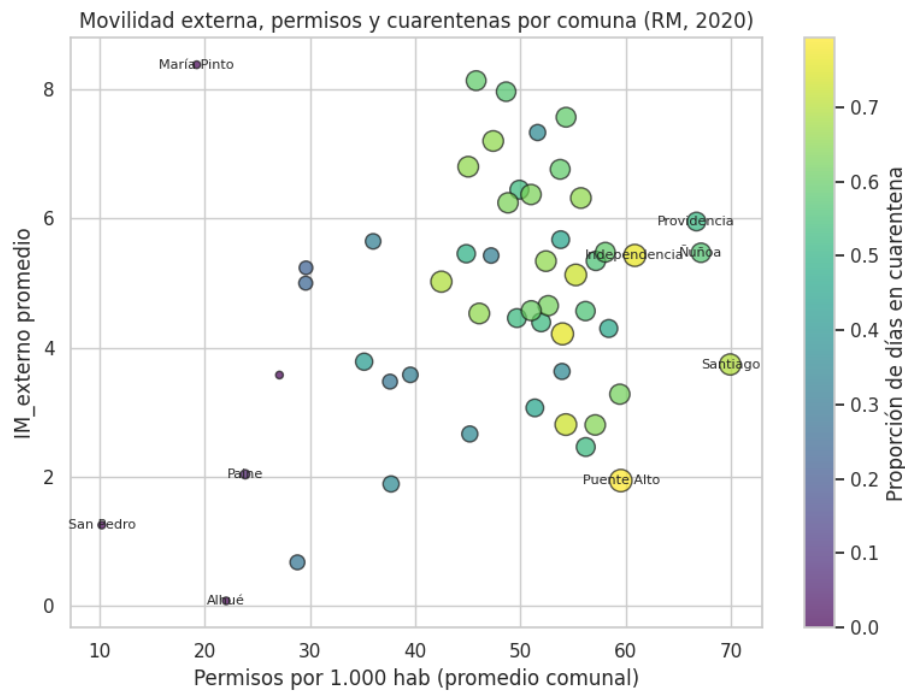


Fuente: Elaboración propia.

Las comunas con mayor movilidad externa promedio (parte superior de la figura) no son necesariamente las que presentan mayor intensidad de permisos ni las que pasan más días en cuarentena. Por ejemplo, María Pinto y El Monte muestran alta movilidad con bajos niveles de permisos y cuarentena, mientras que comunas centrales como Santiago, Nuñoa o Providencia concentran permisos per cápita elevados y una fracción importante del periodo bajo cuarentena, pero con niveles de movilidad externa más moderados.

La siguiente figura resume la heterogeneidad comunal combinando, en un mismo plano, la movilidad externa promedio (IM_externo), la intensidad de permisos por 1.000 habitantes y la fracción de días en cuarentena. En el eje horizontal se ubican las comunas según su nivel promedio de permisos per cápita, mientras que el eje vertical representa la movilidad externa promedio. El tamaño y color de las burbujas indican la proporción de días en que cada comuna estuvo en cuarentena.

Figura 3 – Scatter tipo burbuja por comuna



Fuente: Elaboración propia.

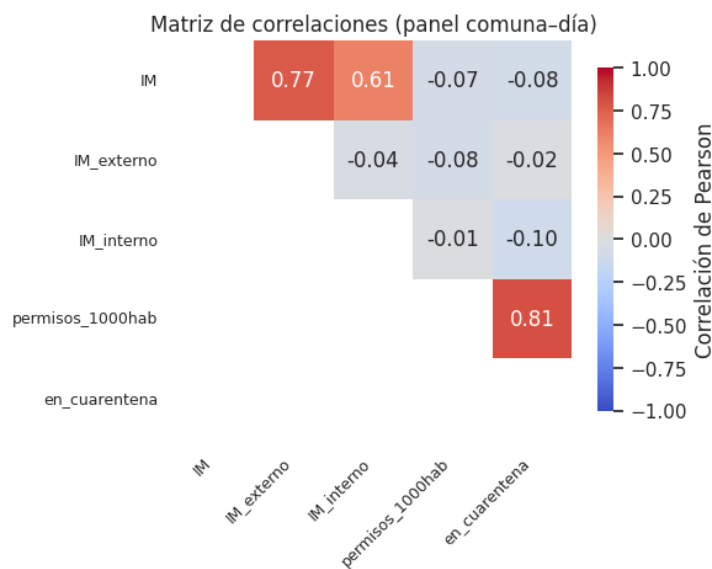
Se observa un grupo de comunas centrales y densas, como Santiago, Ñuñoa, Providencia e Independencia, concentradas en la zona de altos permisos por 1.000 habitantes y niveles intermedios de movilidad externa, con burbujas grandes y claras que reflejan una alta proporción de días en cuarentena (en torno a 0,5–0,75 del periodo). En contraste, comunas como María Pinto, San Pedro o Alhué aparecen con movilidad externa promedio muy baja y con pocos permisos por 1.000 habitantes y prácticamente sin días en cuarentena (burbujas pequeñas y oscuras).

En conjunto, el gráfico refuerza la idea de que la relación entre permisos, cuarentenas y movilidad es marcadamente heterogénea entre comunas y que los mayores volúmenes de permisos y las cuarentenas más prolongadas se concentran en comunas urbanas del eje centro–oriente, mientras que comunas rurales o periféricas muestran configuraciones muy distintas, donde los permisos no necesariamente acompañan una reducción sistemática de la movilidad externa.

Correlaciones entre movilidad, permisos y cuarentenas

Antes de estimar modelos, es conveniente visualizar la relación estadística entre los índices de movilidad, los permisos per cápita y las cuarentenas. Para ello se construye, en primer lugar, una matriz de correlaciones a nivel de todo el panel comuna–día, y en segundo lugar se analiza cómo varía la correlación entre movilidad externa y permisos por comuna.

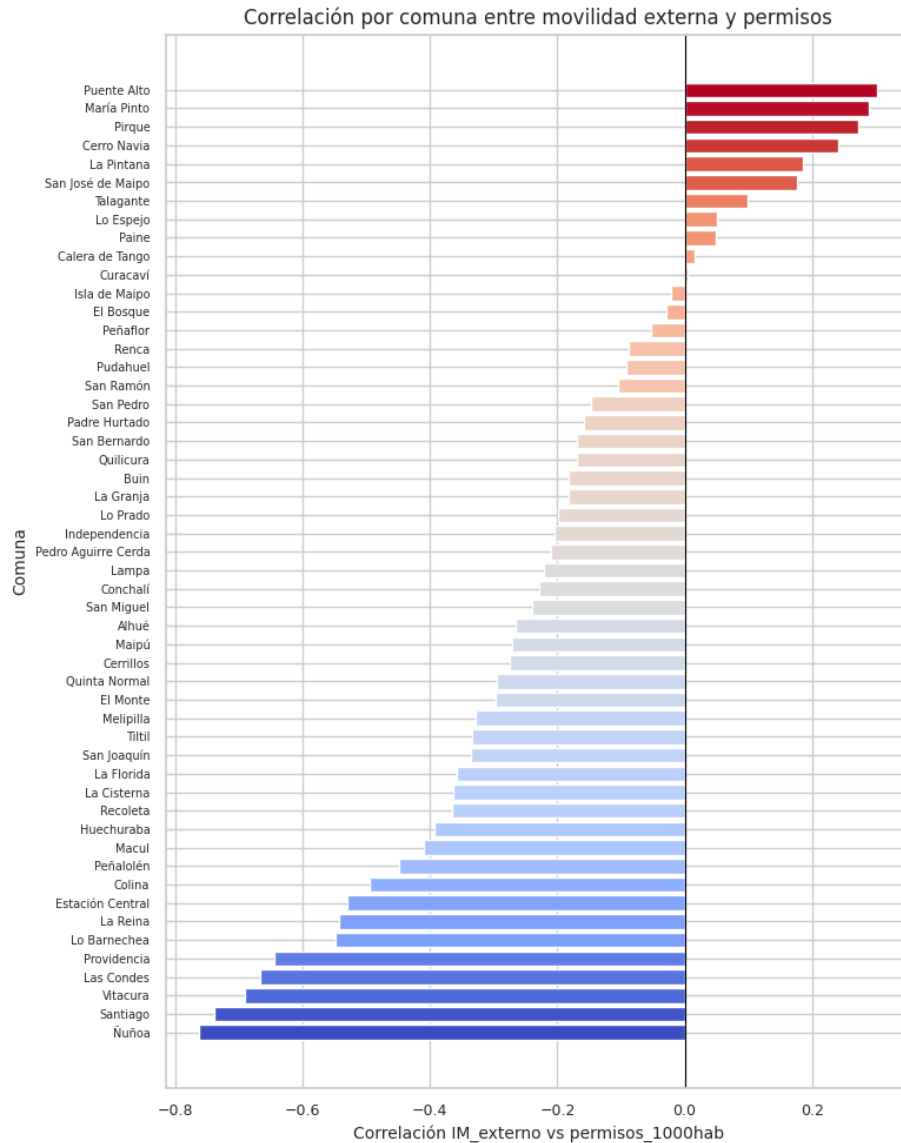
Figura 4: Matriz de correlaciones del panel comuna–día



Fuente: Elaboración propia.

La Figura 4 muestra la matriz de correlaciones de Pearson entre los índices de movilidad (IM, IM_externo, IM_interno), la tasa de permisos por 1.000 habitantes (permisos_1000hab) y el indicador de cuarentena (en_cuarentena), calculada sobre todas las observaciones del panel. Se observa que los tres índices de movilidad presentan correlaciones positivas y elevadas entre sí, cuando aumenta la movilidad total (IM), tienden a aumentar tanto la componente interna como la externa. Por otro lado, las correlaciones entre movilidad y permisos per cápita son bajas en magnitud y ligeramente negativas, mientras que la correlación entre movilidad externa y el indicador de cuarentena es cercana a cero. Por otro lado, permisos per cápita y cuarentena muestran una correlación claramente positiva, consistente con el hecho de que el sistema de permisos se activa y se intensifica precisamente durante los periodos en que las comunas se encuentran bajo restricciones sanitarias.

Figura 5: Correlación por comuna entre movilidad externa y permisos

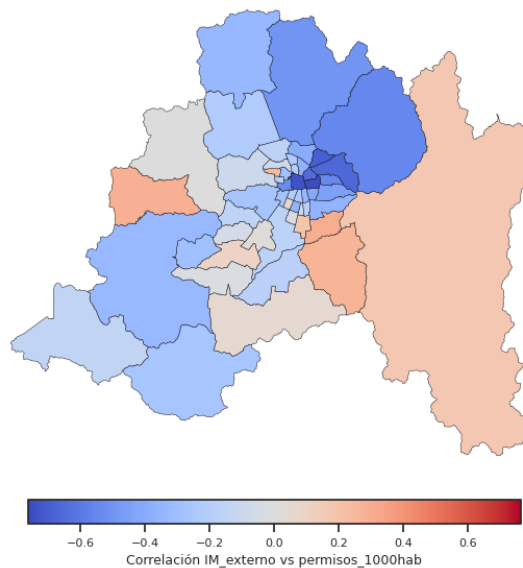


La Figura 5 complementa el análisis anterior mostrando, para cada comuna de la Región Metropolitana, la correlación entre la movilidad externa (IM_externo) y la tasa de permisos por 1.000 habitantes (permisos_1000hab). Las barras horizontales se ordenan desde las correlaciones más negativas a las más positivas, y el color refleja el signo y la magnitud de la correlación. En la distribución se observa que predominan correlaciones negativas para la mayoría de las comunas, los días con mayor tasa de permisos tienden a

coincidir con menor movilidad intercomunal observada. Destacan un conjunto de comunas del eje centro-oriental con Ñuñoa y Santiago como los casos más extremos, seguidas por Vitacura, Las Condes, Providencia, Lo Barnechea y La Reina con correlaciones altas en magnitud y negativas (en torno a $-0,5$ a $-0,8$). Este patrón es consistente con que esas comunas comparten rasgos de “núcleo” económico y alta concentración de empleo y servicios. En el extremo opuesto, un grupo acotado de comunas presenta correlaciones positivas (máximos cercanos a $+0,2/+0,3$, como Puente Alto, María Pinto, Pirque, Cerro Navia y La Pintana), donde permisos y movilidad externa tienden a moverse en el mismo sentido, aunque en menor magnitud, lo que podría reflejar diferencias en patrones de desplazamiento en esas zonas de la región.

En conjunto, la relación permisos–movilidad varía de forma importante entre comunas, lo que anticipa la necesidad de modelar explícitamente esas diferencias (vía efectos fijos por comuna) en el análisis posterior.

Figura 6: Región Metropolitana: correlación entre movilidad externa y permisos por comuna

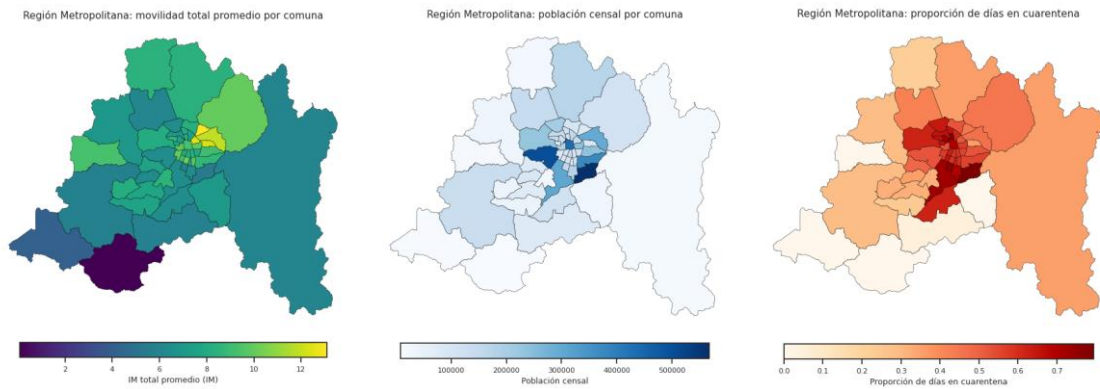


Fuente: Elaboración propia con datos del BCN.

La Figura 6 aporta una visión territorial y muestra que la relación permisos–movilidad no solo varía entre comunas, sino que además presenta un patrón espacial claro. Las correlaciones más negativas (azules más intensos) se concentran en el continuo urbano del centro y centro-oriente, formando un clúster de comunas contiguas asociado al principal núcleo de empleo y servicios de la región. A medida que se avanza hacia comunas más periféricas, el mapa exhibe una transición hacia tonos más neutros (correlaciones cercanas a cero), y aparecen focos acotados de correlación positiva en algunas comunas de la periferia, incluyendo zonas extensas y menos densas donde permisos y movilidad externa tienden a moverse en el mismo sentido, aunque con magnitudes más moderadas. En conjunto, el mapa refuerza que la asociación entre permisos y movilidad externa está estructurada territorialmente y es consistente con diferencias persistentes entre comunas.

Para interpretar mejor estos patrones, a continuación se presentan mapas separados de movilidad externa promedio, población comunal y fracción de días en cuarentena, lo que permite contrastar la distribución espacial de estas variables con el mapa de correlaciones e identificar comunas donde la baja (o alta) asociación entre permisos y movilidad podría estar condicionada por diferencias estructurales del territorio más que por el uso de permisos en sí mismo.

Figura 7: Mapas comparativos de movilidad, población y cuarentenas en la Región Metropolitana



Fuente: Elaboración propia con datos de la BCN.

La Figura 7 resume la heterogeneidad territorial de la Región Metropolitana en tres dimensiones. En primer lugar, la movilidad total promedio (imagen izquierda) se concentra en el continuo urbano central y sus comunas adyacentes, mientras que las comunas más periféricas exhiben niveles promedio menores, consistente con diferencias estructurales de centralidad, conectividad y actividad económica. En segundo lugar, la población censal (imagen central) también se concentra en el núcleo metropolitano, evidenciando que una parte relevante de la variación observada en movilidad ocurre precisamente donde reside (y se desplaza) una mayor proporción de habitantes. Finalmente, la proporción de días en cuarentena (imagen derecha) muestra una mayor cantidad de restricciones en comunas centrales, sugiriendo que la política sanitaria no se distribuyó homogéneamente en el territorio, sino que tendió a focalizarse donde la densidad y la posibilidad de desplazamientos eran mayores. En conjunto, estos tres mapas muestran que el funcionamiento efectivo del sistema de permisos y su relación con la movilidad dependen fuertemente del territorio, lo que justifica el uso de modelos de panel con efectos fijos por comuna en la sección siguiente.

5.2. Modelos de regresión lineal con efectos fijos

El análisis exploratorio anterior muestra que la movilidad externa, los permisos de desplazamiento y las cuarentenas presentan una fuerte variación temporal y heterogeneidad entre comunas. Para cuantificar estas relaciones se estiman modelos de regresión lineal para datos de panel a nivel comuna–día, donde la variable dependiente es el índice de movilidad externa (IM_externo) y los regresores incluyen la tasa de permisos por 1.000 habitantes (permisos_1000hab), un indicador diario de cuarentena (en_cuarentena) y controles de calendario (día de la semana y mes).

Se consideran tres especificaciones:

- Un modelo Pooled OLS sin efectos fijos por comuna, que entrega una referencia “naive” de la asociación promedio en el panel.
- Un modelo de efectos fijos por comuna (FE reducido) que incorpora EntityEffects e incluye cuarentena y controles temporales, capturando heterogeneidad inobservable constante en el tiempo a nivel comunal.
- Un modelo de efectos fijos por comuna (FE completo) que añade permisos_1000hab.

En todas las estimaciones se utilizan errores estándar robustos clusterizados por comuna, para permitir correlación serial y heterocedasticidad dentro de cada unidad territorial.

La Tabla 3 resume las tres especificaciones consideradas para IM_externo (Pooled OLS, FE reducido y FE completo), junto con los principales estadísticos de ajuste y significancia.

Tabla 3: Modelos de regresión lineal de IM_externo con y sin efectos fijos por comuna.

Modelo	Pooled OLS	FE Reducido	FE Full
R ² (overall)	0.144	—	—
R ² (within)	0.423	0.470	0.484
β permisos_1000hab	-0.0075	—	-0.0075
p permisos_1000hab	0.039	—	<0.001
β en_cuarentena	0.8232	-0.5527	-0.1169
p en_cuarentena	0.121	<0.001	0.170
N	10,816	10,816	10,816

Fuente: Elaboración propia

En el primer modelo (Pooled OLS sin efectos fijos por comuna), el Indicador de movilidad Externo se explica por la cantidad de permisos por 1000 habitantes, el indicador de cuarentena y controles de calendario (día de la semana y mes), utilizando errores estándar robustos clusterizados por comuna. Este modelo base presenta un R² overall de 0,144 (y un R² within de 0,423), lo que indica que, si bien el modelo captura parte de la variación temporal dentro de las comunas, una porción importante de la variación total se relaciona con diferencias persistentes entre territorios que el modelo pooled no absorbe. En esta especificación, el coeficiente asociado a permisos_1000hab es negativo y estadísticamente significativo ($\beta = -0,0075$; $p = 0,039$), indicando que, manteniendo constantes cuarentena y calendario, un aumento de 10 permisos por 1.000 habitantes se asocia con una disminución promedio de aproximadamente 0,075 unidades en IM_externo.

Para contrastar lo anterior, se incorporan los modelos con efectos fijos por comuna (EntityEffects), los cuales exploran únicamente la variación dentro de cada comuna en el tiempo y controlan por toda heterogeneidad inobservable constante (por ejemplo, estructura urbana, nivel socioeconómico, conectividad o cantidad de habitantes). En el

modelo FE reducido, que incluye en_cuarentena y controles de calendario pero excluye permisos, el ajuste mejora en términos de variación intra-comunal (R^2 within = 0,470). En esta especificación, en_cuarentena presenta un efecto negativo y estadísticamente significativo ($\beta = -0,5527$; $p < 0,001$), consistente con una caída de la movilidad externa cuando la comuna entra en cuarentena, una vez descontadas las diferencias fijas entre territorios y los patrones sistemáticos de calendario. En magnitud, el coeficiente sugiere una reducción promedio cercana a 0,55 unidades de IM_externo asociada al estado de cuarentena, manteniendo constantes los otros efectos temporales.

Al incorporar permisos_1000hab en el modelo FE completo, el ajuste aumenta levemente en 0,014 (con un R^2 within = 0,484), lo que indica que la variación adicional capturada por permisos dentro de cada comuna aporta información marginal. En este modelo, permisos_1000hab muestra un coeficiente negativo y estadísticamente significativo ($\beta = -0,0075$; $p < 0,001$), reforzando la evidencia de que, dentro de una misma comuna, aumentos en la tasa de permisos se asocian a disminuciones en IM_externo. En cambio, el coeficiente de en_cuarentena se reduce considerablemente en magnitud y pierde significancia estadística ($\beta = -0,1169$; $p = 0,170$). Este cambio entre el modelo reducido y el completo sugiere que ambas variables, permisos y cuarentena, comparten una fracción importante de la variación y que, al incluir ambos simultáneamente, el modelo tiene menor capacidad para atribuir de forma precisa un efecto parcial independiente a la cuarentena.

Este resultado es coherente con la alta correlación observada entre ambas variables (corr overall $\approx 0,819$; corr within $\approx 0,827$), lo cual indica que, especialmente dentro de cada comuna, los permisos tienden a aumentar precisamente en períodos de cuarentena.

Tabla 4: Correlación (overall y within) y factores de inflación de varianza (VIF) de los regresores principales.

Variab les	Corr overall	Corr within	VIF
Permisos	0.818	0.826	3.03
Cuarentena			3.03

Fuente: Elaboración propia.

Aunque los factores de inflación de varianza sugieren una multicolinealidad moderada ($VIF \approx 3,0$ en ambos regresores), la alta correlación temporal implica que la interpretación de los coeficientes debe realizarse con precaución. En particular, la estimación asociada a `permisos_1000hab` se mantiene estable y con mayor precisión, mientras que el efecto atribuido a `en_cuarentena` resulta sensible a la inclusión de `permisos_1000hab` en la especificación, reduciéndose en magnitud y perdiendo significancia estadística en el modelo completo.

En conjunto, los modelos con efectos fijos respaldan que las variaciones diarias en permisos y restricciones están estrechamente vinculadas a cambios en la movilidad externa, pero también evidencian que separar empíricamente los efectos de cuarentena y permisos es difícil dado su fuerte solapamiento en el tiempo.

Como evidencia complementaria, se evaluó la contribución incremental de `permisos_1000hab` mediante un test de Wald sobre $H_0: \beta_{\text{permisos_1000hab}} = 0$ en el modelo con efectos fijos y errores estándar clusterizados por comuna. El contraste rechaza la hipótesis nula ($\chi^2 = 25.32$; $p = 4.86 \times 10^{-7}$), confirmando que la inclusión de permisos aporta información estadísticamente significativa respecto del modelo reducido.

5.3. Random Forest como modelo no lineal de referencia

Random Forest es un método de aprendizaje supervisado basado en árboles de decisión que combina múltiples árboles entrenados sobre subconjuntos aleatorios de datos y variables. En este estudio se utiliza para evaluar si la información disponible permite predecir variaciones de IM_externo con mayor flexibilidad funcional y para explorar la relevancia relativa de las variables explicativas, aportando evidencia complementaria a los resultados de los modelos de efectos fijos.

El modelo se entrena sobre 8.684 observaciones (conjunto de entrenamiento) y se evalúa en 2.132 observaciones correspondientes al tramo final del período (conjunto de prueba) a partir del 10-08-2020, definido mediante un corte en el percentil 80 de la variable fecha, lo que evita incorporar información futura en el entrenamiento. Como predictores se utilizan la tasa de permisos por 1.000 habitantes (permisos_1000hab), el indicador de cuarentena (en_cuarentena) y variables categóricas para comuna (cid), día de la semana (dow) y mes (mes), incorporadas mediante codificación one-hot en un pipeline. El Random Forest se ajusta con 500 árboles y una semilla aleatoria fija (random_state = 42) para asegurar reproducibilidad.

Se probó una variante regularizada del Random Forest para reducir sobreajuste; si bien disminuyó el ajuste en entrenamiento, el desempeño en prueba empeoró (R^2 test 0,452 vs 0,646). Por este motivo, se reporta como modelo principal la configuración base.

Tabla 5. Desempeño predictivo del Random Forest (conjunto de entrenamiento y prueba).

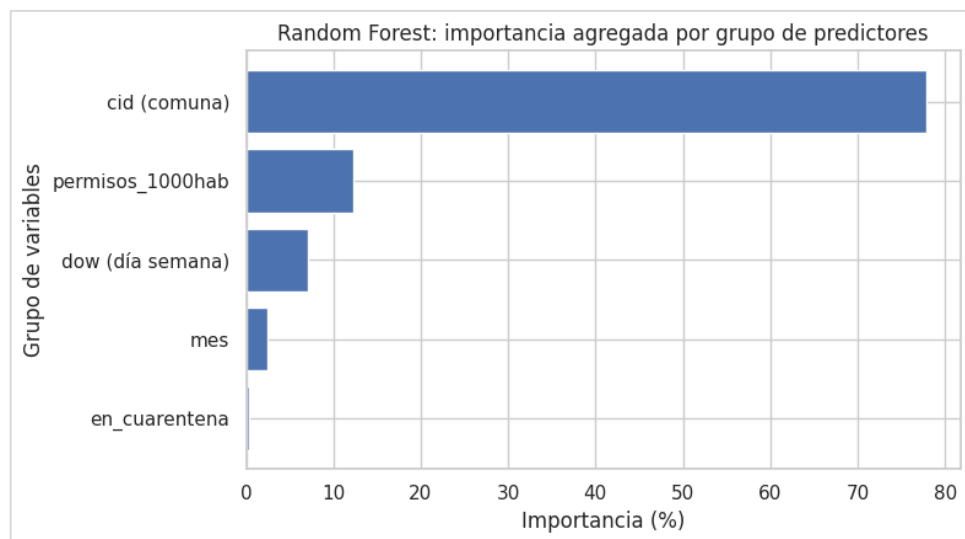
Conjunto	R ²	RMSE	MAE	N
Entrenamiento	0,971	0,367	0,183	8.684
Prueba	0,646	1.469	1,072	2.132

Fuente: Elaboración propia.

La Tabla 5 muestra que el modelo logra un ajuste muy alto en el conjunto de entrenamiento ($R^2 = 0,971$; $RMSE = 0,367$; $MAE = 0,183$) y un desempeño moderado en el conjunto de prueba ($R^2 = 0,646$; $RMSE = 1,469$; $MAE = 1,072$). Es decir, el Random Forest explica aproximadamente un 65% de la variación de $IM_externo$ en el tramo final del período; sin embargo, el deterioro del desempeño al pasar de entrenamiento a prueba indica que el modelo captura patrones específicos del período de entrenamiento que no se replican completamente en las fechas posteriores. Este comportamiento podría indicar que una fracción de la movilidad diaria en el período de prueba responde a factores no observados no representados por las variables disponibles.

Para interpretar el modelo se analizan, en primer lugar, se analizan las importancias de variables del Random Forest, calculadas a partir de la contribución de cada predictor a la reducción del error a lo largo de los árboles y agregadas por grupos (comuna, calendario y variables de política). Este análisis permite identificar qué fuentes de información explican en mayor medida las predicciones de $IM_externo$, con un objetivo descriptivo y sin interpretación causal directa.

Figura 8: Mean Decrease in Impurity por grupo de predictores.



Fuente: Elaboración propia.

La Figura 8 muestra la importancia agregada por grupos de predictores en el Random Forest para explicar IM_externo. Las variables asociadas a la comuna (cid) concentran la mayor parte de la contribución predictiva (0,779 del total), lo que sugiere que el modelo se apoya principalmente en diferencias persistentes entre comunas para ajustar la movilidad externa. En segundo lugar aparece la intensidad de permisos (permisos_1000hab), con una contribución relevante (0,123), seguida por los efectos de calendario, particularmente el día de la semana (0,072) y, en menor medida, el mes (0,024). En contraste, el indicador en_cuarentena exhibe una importancia marginal (0,003), lo que indica que, dado el resto de las variables incluidas, aporta poca información adicional para la predicción.

Este resultado es consistente con la evidencia de los modelos con efectos fijos, donde el coeficiente de en_cuarentena pierde significancia al incorporar permisos_1000hab, lo que es coherente con la fuerte relación entre ambas medidas.

La importancia basada en reducción de impureza (MDI) puede favorecer predictores con mayor granularidad, por lo que se calculó adicionalmente importancia por permutación en el conjunto de prueba, medida como la caída promedio en R^2 al permutar cada variable.

Tabla 6: Importancia por permutación en conjunto de prueba (ΔR^2).

variable	perm_importance_mean	perm_importance_std
cid	1.426	0.026
dow	0.148	0.004
en_cuarentena	0.011	0.001
permisos_1000hab	0.006	0.004
mes	-0.028	0.004

Fuente: Elaboración propia.

Los resultados confirman la importancia de la heterogeneidad territorial. Permutar comunas (cid) reduce el desempeño en 1,43 puntos de R^2 en promedio, muy por encima del resto de predictores. En segundo término, el día de la semana presenta una

contribución relevante ($\Delta R^2 \approx 0,148$). En cambio, las variables de política muestran aportes marginales en términos predictivos una vez consideradas las diferencias por comuna y calendario ($\Delta R^2 \approx 0,011$ para `en_cuarentena` y $0,006$ para `permisos_1000hab`), lo que es consistente con la variación simultánea entre ambas medidas. Finalmente, el mes no aporta en el tramo de prueba (ΔR^2 ligeramente negativo), sugiriendo que su información es redundante o poco estable respecto de otros controles temporales.

5.4.XGBoost como modelo de gradiente boosting

Como complemento al Random Forest, se aplica un modelo XGBoost (gradient boosting con árboles), que también permite capturar relaciones no lineales e interacciones, pero construye los árboles de forma secuencial, de forma que cada árbol nuevo corrige parte de los errores del conjunto anterior. A diferencia del Random Forest, XGBoost incorpora opciones de regularización y control de complejidad (por ejemplo, limitar la profundidad de los árboles o penalizar modelos demasiado complejos), lo que puede ayudar a manejar el sobreajuste. En este trabajo, se utiliza XGBoost como un tercer enfoque para comparar su desempeño con la regresión con efectos fijos y el Random Forest, y para evaluar si los resultados asociados a `permisos_1000hab` se mantiene utilizando un modelo más flexible.

El modelo se entrena sobre el mismo panel comuna-día y con el mismo conjunto de predictores utilizado en Random Forest: `permisos_1000hab`, `en_cuarentena` y variables categóricas para comuna (`cid`), día de la semana (`dow`) y mes (`mes`), incorporadas mediante codificación one-hot. Se mantiene la partición entre entrenamiento y prueba, de modo que el modelo se ajusta con 8.684 observaciones iniciales y se evalúa en 2.132 observaciones correspondientes al tramo final del período. Para seleccionar la complejidad del modelo se entrenó inicialmente con un máximo de 2.000 iteraciones y early stopping, identificándose un número óptimo de árboles. Luego el modelo final se reentrenó sobre todo el conjunto de entrenamiento usando ese número de iteraciones y un conjunto acotado de hiperparámetros (por ejemplo, `learning_rate = 0,05`, `max_depth = 6`,

subsample = 0,8, colsample_bytree = 0,8 y regularización L2), fijando random_state = 42 para asegurar reproducibilidad.

Tabla 7: Desempeño predictivo de XGBoost (conjunto de entrenamiento y prueba).

Conjunto	R²	RMSE	MAE	N
Entrenamiento	0,94	0,48	0,30	8684
Prueba	0,67	1,41	1,04	2132

Fuente: Elaboración propia.

La Tabla 7 muestra que el modelo alcanza un ajuste alto en el conjunto de entrenamiento ($R^2 = 0,94$; $RMSE = 0,48$; $MAE = 0,30$) y un mejor desempeño en el conjunto de prueba ($R^2 = 0,67$; $RMSE = 1,41$; $MAE = 1,04$). En otras palabras, XGBoost explica aproximadamente un 67% de la variación diaria de $IM_externo$ en el tramo final del período. Bajo el mismo esquema de evaluación, este desempeño supera al Random Forest base ($R^2 \text{ test} = 0,646$; $RMSE = 1,469$; $MAE = 1,072$), lo que sugiere que el enfoque de boosting logra capturar con mayor precisión patrones relevantes para generalizar fuera de muestra. No obstante, la diferencia entre entrenamiento y prueba indica que el modelo aún tiene cierta dependencia de patrones del período de entrenamiento y que una fracción de la variabilidad de $IM_externo$ permanece asociada a factores no observados o no incorporados en los predictores.

Al igual que en el caso del Random Forest, se analiza la importancia de las variables a partir de las importancias internas de XGBoost, agregando las dummies en cinco categorías: comuna (cid), permisos_1000hab, día de la semana (dow), mes (mes) y en_cuarentena. La Tabla 10 resume las importancias agrupadas.

Tabla 8: Importancia relativa por grupos de variables en XGBoost.

Grupo	Importancia relativa
cid (comuna)	0,9806
dow (día semana)	0,0092
mes	0,0068
permisos_1000hab	0,0022
en_cuarentena	0,0012

Fuente: Elaboración propia.

La Tabla 8 muestra la importancia de predictores en XGBoost (medida como gain). Los resultados muestran que las variables asociadas a la comuna (cid) concentran prácticamente toda la capacidad explicativa del modelo (98,1%), mientras que los efectos de calendario tienen un peso muy reducido (dow: 0,9%; mes: 0,7%). Por otro lado, las variables de política (permisos_1000hab y en_cuarentena) presentan importancias marginales (0,22% y 0,12%, respectivamente) lo que indica que, en el contexto de este modelo, su contribución incremental a la reducción del error es pequeña cuando se incorporan las diferencias entre comunas y el calendario.

En conjunto, estos resultados son consistentes con la evidencia previa de la regresión con efectos fijos y del Random Forest. La mayor parte de la capacidad predictiva se asocia a diferencias estructurales entre comunas y a patrones de calendario, mientras que los permisos de desplazamiento aportan información adicional pero acotada. Asimismo, el indicador de cuarentena muestra una contribución marginal reducida una vez que se controlan simultáneamente los demás factores, en un contexto donde permisos y cuarentena presentan una alta co-variación temporal.

6. Conclusiones

Este trabajo evaluó cómo se relacionan los permisos de desplazamiento por 1.000 habitantes y las cuarentenas locales con la movilidad externa observada en la Región Metropolitana durante 2020, usando un panel comuna-día. Para obtener una lectura completa se combinaron modelos de panel (pooled y efectos fijos por comuna) con modelos supervisados basados en árboles (Random Forest y XGBoost) bajo un esquema de evaluación temporal.

Los resultados de las regresiones muestran que, al comparar cada comuna consigo misma en el tiempo (efectos fijos), un mayor nivel de permisos por 1000 habitantes se asocia con una menor movilidad externa ($\beta = -0,0075$; significativo). En cambio, el efecto parcial de cuarentenas es más inestable: Su magnitud y significancia dependen de si el modelo incluye o no permisos, lo que es consistente con la fuerte relación entre ambas variables (correlación alta y VIF moderado). En términos prácticos, la evidencia indica que los permisos capturan parte de la variación de movilidad asociada a los períodos de restricciones, y que separar completamente “cuarentena” versus “permisos” es difícil cuando ambas medidas se mueven casi en simultáneo.

Desde el punto de vista predictivo, los modelos de árboles confirman que la mayor parte de la explicación de la movilidad está asociada a diferencias estructurales entre comunas y, en segundo orden, a patrones de calendario. En el corte temporal de prueba, XGBoost entrega el mejor desempeño ($R^2 = 0,673$; RMSE = 1,413; MAE = 1,037), superando a Random Forest. En conjunto, se puede concluir que la movilidad externa en 2020 está fuertemente condicionada por el componente territorial, mientras que permisos y cuarentenas aportan información más acotada.

En resumen, la evidencia del panel indica que, una vez controladas las diferencias fijas entre comunas y los efectos de calendario, la tasa de permisos por 1.000 habitantes mantiene una asociación negativa y estadísticamente significativa con la movilidad externa, sin embargo la movilidad externa en 2020 está fuertemente condicionada por el componente territorial, mientras que permisos y cuarentenas aportan información más acotada. En segundo lugar, el efecto asociado a la variable de cuarentena no muestra la misma estabilidad, ya que su magnitud y significancia cambian al incorporar permisos en la especificación. Esto no invalida el rol de las cuarentenas, sino que refleja un problema práctico por la naturaleza de estos datos, dado que permisos y cuarentena presentan una evolución altamente sincronizada, lo que limita la capacidad de atribuir variaciones de movilidad de manera separada a cada componente cuando ambos se mueven en paralelo.

Como próximos pasos, este trabajo deja un marco operativo que puede reutilizarse ante escenarios futuros en que sea necesario monitorear y gestionar restricciones de movilidad. En particular, la construcción del panel comuna-día y la comparación sistemática entre enfoques econométricos y predictivos permiten identificar rápidamente qué parte de los cambios en movilidad responde a patrones territoriales persistentes y qué parte se asocia a variaciones en instrumentos de control. En ese sentido, el enfoque propuesto puede servir como base para tableros de seguimiento y análisis orientados a apoyar la toma de decisiones de autoridades sanitarias y de transporte en eventuales nuevas emergencias, sin requerir rediseñar desde cero la integración de fuentes ni la lógica de evaluación, y facilitando una interpretación consistente y comparable de la movilidad observada bajo distintos regímenes de restricciones.

Bibliografía

- Pappalardo, L., Cornacchia, G., Navarro, V., Bravo, L., & Ferres, L. (2023). A dataset to assess mobility changes in Chile following local quarantines. *Scientific Data*, 10(1), 6. <https://doi.org/10.1038/s41597-022-01893-3>
- Araya-Moreno, J. (2022). Pandemic documents: The case of mobility permits during lockdown in Chile. *Anthropologica*, 63(2). <https://doi.org/10.18357/anthropologica63220211031>
- Oliver, N., et al. (2020). Mobile phone data for informing public health actions across the COVID-19 pandemic life cycle. *Science Advances*, 6(23), eabc0764. <https://doi.org/10.1126/sciadv.abc0764>
- Cameron, A. C., & Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, 50(2), 317–372. <https://doi.org/10.3368/jhr.50.2.317>
- Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data* (2nd ed.). The MIT Press. <https://mitpress.mit.edu/9780262232586/econometric-analysis-of-cross-section-and-panel-data/>
- Mobility Index for Local Quarantines in Chile. Datos de Movilidad y Cuarentenas. <https://doi.org/10.6084/m9.figshare.c.5214272>
- Censo 2024. Base de datos de personas censadas. <https://censo2024.ine.gob.cl/resultados/>
- Biblioteca del Congreso Nacional. Datos de Mapas vectoriales de las comunas de Chile. https://www.bcn.cl/siit/mapas_vectoriales/index_html