



puntos de referencia

CENTRO
DE ESTUDIOS
PÚBLICOS

EDICIÓN DIGITAL
N° 718, DICIEMBRE 2024

ECONOMÍA Y POLÍTICAS PÚBLICAS

La distribución comunal del delito en Chile.

Un análisis exploratorio en base a ENUSC 2023

CÉSAR GAMARRA, ALDO MASCAREÑO, ARIEL CARDEIRO Y JUAN ROZAS

C22, Aprender de Chile con métodos digitales
<https://c22cepchile.cl/>



RESUMEN

- El análisis territorial del delito es un instrumento fundamental para la toma de decisiones de política pública en relación con la seguridad ciudadana. La Encuesta ENUSC 2023 ha incluido por primera vez datos a nivel comunal para 136 comunas de las 346 de Chile.
- Por medio de distintas técnicas estadísticas avanzadas, en este artículo exploramos esos datos para identificar líneas de análisis e investigación relevantes para enfrentar la actual situación de criminalidad en el país.
- Los resultados muestran que las zonas de mayor riesgo, con énfasis en distintos tipos de delito, son el sector poniente de la Región Metropolitana, algunas comunas de la zona norte y otras en las zonas centro sur y sur.
- El artículo concluye que la política de seguridad pública debe realizar un acto de balance entre dos tipos de equilibrio: por un lado, el equilibrio entre contención del delito e intervención en los entornos que lo producen y, por otro, el de la relación entre focalización en zonas de riesgo y la protección general de la población en el Estado de derecho.

Palabras clave: delitos, delincuencia, distribución comunal del delito, macrozonas, ENUSC 2023

CÉSAR GAMARRA es investigador asistente del Centro de Estudios Públicos.

ALDO MASCAREÑO es investigador senior del Centro de Estudios Públicos.

ARIEL CARDEIRO es pasante del Centro de Estudios Públicos.

JUAN ROZAS es investigador asistente del Centro de Estudios Públicos.

Los autores agradecen las sugerencias y comentarios metodológicos provistos por Ariadna Chuaqui, Ignacio Díaz y Mauricio Salgado.

1.

INTRODUCCIÓN

La comprensión del fenómeno delictual es vital para la formulación de políticas públicas adecuadas que controlen sus impactos. La utilización de información basada únicamente en registros policiales posee severas limitaciones, principalmente debido a la gran cantidad de delitos que las víctimas deciden no reportar (Biderman y Reiss 1967; Skogan 1977; Brunton-Smith et al. 2024). En ese sentido, numerosa evidencia anuncia la utilidad de las encuestas de victimización para complementar y ampliar el conocimiento sobre las actividades delictuales, proveyendo evidencia accesible e iluminadora sobre la naturaleza de la distribución e incidencia del delito (Lohr 2019; Xie y Baumer 2019).

Utilizando la Encuesta Nacional Urbana de Seguridad Ciudadana (ENUSC 2023), que contiene información sobre los delitos reportados y no reportados, el objetivo de este artículo es explorar la distribución del delito en Chile al nivel más granular posible con los datos indicados: la comuna. Buscamos analizar si esta distribución comunal de determinados delitos correlaciona con factores como la percepción de inseguridad, la evaluación de la policía y otras variables construidas por medio de análisis factorial. Para esto empleamos los datos de última versión disponible de la ENUSC, aplicada a 49.813 viviendas en 136 de las 346 comunas del país, cuyas innovaciones metodológicas ya permiten la desagregación de información a nivel de municipios (ENUSC 2023).

Tratándose de un estudio exploratorio, ponemos énfasis en la aplicación de varios procedimientos metodológicos y técnicas estadísticas. Inicialmente, calculamos la proporción de hogares victimizados por 16 categorías de delitos. Estimamos esta proporción con un promedio *bootstrap*. Luego, empleamos un análisis factorial que nos permite establecer cómo las categorías de delito previamente calculadas correlacionan entre sí formando dimensiones de delitos. Seleccionamos los cuatro primeros factores que capturan más del 50% de la variabilidad de los datos a nivel comunal, y con ellos generamos clústeres de comunas por medio del algoritmo no supervisado de *k-media* (Schwarz 2008; Mackles 2012). De esta forma, podemos observar una distribución heterogénea de delitos en comunas particulares, es decir, cómo determinados delitos —contenidos en una dimensión de delitos— ocurren con mayor incidencia en un clúster (grupo de comunas). Finalmente, realizamos una caracterización de los clústeres en términos de varios índices generados (confianza en instituciones, evaluación de la labor de la municipalidad, percepción del aumento de la delincuencia) y concluimos indicando cómo las dimensiones de delitos correlacionan con algunos de estos índices. También presentamos regresiones lineales multivariadas para establecer si los índices generados correlacionan con los cuatro factores construidos.

En términos generales, los resultados muestran una distribución desigual del delito a nivel comunal que correlaciona con las percepciones de inseguridad, si bien estas son generalizadas a nivel país. En base a estos análisis identificamos que una zona especialmente crítica es el sector poniente de la Región

Metropolitana; otros territorios críticos son comunas particulares de la zona norte del país, además de algunas de la zona centro sur y sur. Los detalles de estas especificaciones se discuten en los resultados del estudio y sus conclusiones.

El artículo inicia con una descripción de los datos y métodos empleados. Continúa con los resultados y el análisis de ellos bajo los distintos procedimientos indicados. Luego elaboramos algunas líneas de exploración y profundización derivadas de los datos. Finalmente, presentamos conclusiones sustantivas y sugerencias sobre el tema de investigación.

2.

DATOS Y MÉTODO

La base de datos empleada es ENUSC (2023). Ella posee información para 49.813 hogares en 136 de las 346 comunas de Chile. Para este estudio, la unidad de análisis es la comuna, lo que resulta en que la máxima cantidad de observaciones posible sea 136, aquellas en que residen los encuestados (victimizados).¹

Distinguimos entre delitos violentos y no violentos (ENUSC 2023). Los delitos violentos están compuestos por: robo con violencia, agresiones, robo por sorpresa, amenazas, extorsiones, robo con fuerza de la vivienda y robo de vehículos. Mientras que los delitos no violentos están compuestos por: robo desde vehículos, hurtos, fraude, estafas, sobornos, vandalismo en la vivienda, vandalismo en vehículo, *hacking*, suplantación de identidad, acoso sexual, *bullying* e infecciones con software maliciosos (virus). Empleamos la probabilidad de que un hogar sea victimizado por alguno de los delitos mencionados para un análisis factorial exploratorio.

Además de la victimización específica por cada delito, utilizamos tres conjuntos de variables adicionales. El primero hace referencia a índices de victimización, revictimización y subreporte a nivel agregado. Identificamos la proporción de hogares victimizados por algún delito violento (victimización agregada), o la proporción —dentro del subconjunto que sí fue victimizado— de revictimización por delito violento (revictimización agregada), y si denunció alguno de los delitos violentos que sufrió (denuncia agregada). Esto se realiza también en delitos no violentos.

El segundo conjunto de variables forma índices generados a través de la suma o promedio de diversos ítems o la proporción de alguna variable. Luego de la generación de este valor a nivel hogar, se promedia el indicador a nivel comunal. Los índices son los siguientes:

¹ Existe, sin embargo, una alta proporción de hogares que sufre los delitos en la misma comuna en la que reside. Esta proporción es más baja para hogares con mayor nivel socioeconómico (i.e., hogares con mayor nivel socioeconómico, probablemente habitantes de comunas con mayores ingresos, poseen una menor probabilidad de sufrir el delito en la comuna que residen). Esta variabilidad, también, depende del delito sufrido. Más adelante se analiza con más detalle este fenómeno.

- Percepción de incremento del delito. Contiene 3 variables a nivel país, comunal y de barrio (porcentaje de personas que anuncia que la delincuencia aumentó).
- Inseguridad en lugares públicos. Suma de lugares públicos en los que el encuestado declara sentirse inseguro o muy inseguro (mínimo de 0 lugares, máximo posible de 16 lugares).
- Presencia de desórdenes. Suma de desórdenes que el encuestado declara que ocurren siempre o casi siempre en su barrio (mínimo de 0, máximo de 8).
- Incivildades. Suma de incivildades que el encuestado declara que ocurren siempre o casi siempre en su barrio (mínimo de 0, máximo 7).
- Evaluación de la policía. Evaluación del trabajo de Carabineros (1 = Muy malo, 5 = Muy bueno) ($\alpha = 0.89$). Se considera el promedio entre 6 evaluaciones.
- Evaluación. Evaluación del municipio en términos de seguridad ciudadana (1 = Muy malo, 5 = Muy bueno) ($\alpha = 0.85$). Se considera el promedio entre 5 evaluaciones.
- Número de medidas adoptadas. Número de acciones adoptadas por el hogar para combatir la delincuencia (9 posibles alternativas).
- Confianza en instituciones. Porcentaje que declara tener Mucha + Bastante confianza en Carabineros, PDI y Fiscalía o Ministerio Público entre quienes conocen la labor de estas instituciones en materia de seguridad ciudadana.
- Apoyo ciudadano entre vecinos. Contiene 5 ítems que miden la probabilidad de que exista apoyo entre vecinos en términos de seguridad ciudadana y actividades conjuntas ($\alpha = 0.66$). Se considera el promedio por hogar (las acciones mencionadas son calificadas entre 1 = Muy probable o 4 = Muy improbable).
- Fuente de información. Si la fuente en la que está basada su percepción de la delincuencia es la experiencia propia o personal (porcentaje de personas que afirma dicha categoría).
- Pronóstico de delitos. Variable dicotómica si el encuestado declara que podría ser víctima de algunos de los delitos consultados en los próximos 12 meses. Se clasificó de la forma:
 - Violentos. Robo en su vivienda, robo o hurto de su vehículo o *portonazo*, robo o asalto con violencia (incluye ‘cogoteo’, robo por sorpresa o lanzazo), agresiones físicas o lesiones y Amenazas o extorsiones (5 delitos).
 - No violentos. Robo o hurto de algún objeto dejado dentro del vehículo o parte de él, vandalismo o daño a su vivienda o vehículo, hurto, delitos económicos como fraude o estafa, delitos cibernéticos, acoso callejero o sexual (6 delitos).

Como tercer conjunto, disponemos de variables de control que son calculadas como porcentajes a nivel comunal. Las variables sociodemográficas son: porcentaje de encuestados con educación superior, extranjeros, GSE alto y bajo, pertenecientes a pueblos indígenas, mujeres, hombres y por tramo etario (4 categorías). Estas variables sociodemográficas son respondidas por todos los miembros del hogar —a diferencia del resto de variables, que únicamente las responde una persona por hogar (*kish*).

Para procesar los datos procedemos con una metodología dividida en tres partes. En primer lugar, presentamos estadística descriptiva (en la forma de hechos estilizados) utilizando los índices generados y analizando la correlación que poseen entre ellos y con respecto a otras variables que muestren victimización o subreporte. Para ello se utilizan métodos lineales bivariados, empleando tanto la pendiente de estimación o el porcentaje de la varianza que la variable independiente captura por sí sola. Las variables de victimización y subreporte —ya sea a nivel agregado, e.g., índice de victimización si es que el hogar fue víctima de algún delito violento o no violento, o de forma desagregada, para cada delito en específico— son estimadas a nivel comunal utilizando una metodología *bootstrap*. Se sigue lo realizado por Gamarra et al. (2024) con 40 repeticiones para muestras de 5.000 observaciones (con reemplazo).²

En segundo lugar, focalizados en el propósito de hacer una caracterización de las formas de violencia delictual que enfrenta cada comuna, realizamos un análisis factorial exploratorio utilizando componentes principales para identificar si existen factores que asocien determinadas categorías de delito, esto es, si es posible generar un factor o dimensión que contenga varias categorías de delito que poseen una alta correlación en su distribución comunal. De esta forma es posible reducir la dimensionalidad del análisis. Evaluamos si dicha información puede ser absorbida por dimensiones que contengan agrupaciones de delitos que varíen de forma similar, en vez de analizar delito a delito. Debido a que el propósito final del análisis factorial es introducir las dimensiones predichas en un análisis de clústeres, aplicamos el método de rotación ortogonal para forzar a que la correlación entre dimensiones sea nula, de forma que cada dimensión posea la mayor cantidad de información única.

Primero, aplicamos el análisis factorial sobre 16 categorías de delito.³ Estas variables representan la proporción de hogares victimizados por cada delito para cada comuna estimada a través de *bootstrap*. La Tabla A1 presenta estadística descriptiva de los delitos a nivel comunal. Segundo, empleamos el método de rotación ortogonal *varimax* para que la correlación entre factores sea nula. Del total de factores probados, 5 poseen *eigenvalues* superiores a 1 (criterio Kaiser). De esos 5, decidimos utilizar 4 factores debido a la proporción de varianza que explican.⁴ La fracción de la varianza total explicada por cada factor se ilustra en la Tabla A2. Tercero, a partir de la correlación entre las variables (delitos) y los 4 factores es posible esclarecer qué variables se encuentran contenidas dentro de cada factor. Analizamos así la composición de la sustracción de dimensiones del análisis factorial. Esta información se presenta en la Tabla A2. En base a ella, con criterios estándar,⁵ desarrollamos la siguiente clasificación:

² Se considera el promedio de estas estimaciones (*bootstrapped mean*). Las variables que fueron presentadas en Gamarra et al. (2024) —victimización y denuncia agregada por delitos violentos y no violentos— son utilizadas en este análisis con los mismos valores.

³ Por motivos de eficiencia en la generación de los clústeres, se decidió sustraer los siguientes delitos del análisis: delitos cibernéticos y acoso callejero o sexual. El procedimiento es presentado de manera secuencial en Anexo 1.

⁴ La consideración de esos factores responde a que cada factor represente al menos un 8% de la varianza total. El factor 4 representa 8.3% de la varianza. Un criterio más elevado limitaba la realización de los clústeres al poseer menos variables para la agrupación (en general, lo ideal sería poseer al menos 4 variables para la agrupación, como es en nuestro caso).

⁵ Se posee el criterio de que el *factor loading* sea superior a 0.39. Sin embargo, para el caso del delito de fraude bancario se vislumbra que cumple dicho criterio en dos dimensiones. Se interpreta que dicho delito está contenido en la dimensión que mayor *factor loading* posee.

- El **factor 1** está principalmente compuesto por: robo con violencia, robo por sorpresa, robo de vehículos, robo desde vehículos y hurtos. Este factor será denominado *robo con violencia*.
- El **factor 2** está principalmente compuesto por agresiones, amenazas, vandalismo de vehículos y vandalismo en la vivienda. Este factor será denominado *agresión y vandalismo*.
- El **factor 3** está compuesto esencialmente por fraude bancario y estada. Este factor será denominado *delitos económicos*.
- El **factor 4** está compuesto principalmente por extorsión y robos en la vivienda. Este factor será denominado: *extorsión y robo en la vivienda*.

Cuarto, con estos factores (robo con violencia, agresiones y vandalismo, delitos económicos y extorsión, y robos en la vivienda) procedemos a la generación de clústeres. Para ello, es necesario predecir cada factor a nivel de la unidad de análisis (comuna). Para la predicción utilizamos el método de *regression scoring*⁶, y para la clusterización empleamos el algoritmo no supervisado de *k-medias*.⁷

La idea subyacente del análisis de clústeres es generar agrupaciones entre comunas en base a las variables utilizadas de forma que se maximice la diferencia entre grupos y minimice la variabilidad dentro de cada grupo (Hastie et al. 2009; Kaufman y Rousseeuw 2009). La utilización del análisis factorial como insumo para la construcción de los clústeres no solamente disminuye la dimensionalidad del cálculo (reducción de 16 a 4 variables), sino que facilita la interpretabilidad de los resultados.

El último paso consiste la definición del número óptimo de agrupaciones (clústeres). Para obtener el número adecuado de clústeres empleamos el método de suma de cuadrados (y su logaritmo), el coeficiente η^2 (cuya interpretación es similar a la del R^2 de los modelos de regresión) y la reducción proporcional del error (PRE) (Mackles 2012). En Anexo 1 presentamos más detalle sobre la clusterización. La Figura A1 muestra los 4 estadísticos mencionados previamente para las posibles 20 primeras agrupaciones de variables. En base a los resultados, decidimos establecer $k = 6$ como el número óptimo. Es decir, obtenemos 6 grupos del total de 134 comunas con información completa para el análisis.

Finalmente, presentamos regresiones lineales multivariadas para establecer si los índices generados pudiesen correlacionar con los 4 factores predichos del análisis factorial. Estos resultados pretenden dar luces sobre una posible relación entre la percepción y la experiencia específica de ciertos tipos (dimensiones) de delitos. Para ello, estimamos la ecuación [1]:

$$Y_c = \alpha + \beta_1 F_1 + \beta_2 F_2 + \beta_3 F_3 + \beta_4 F_4 + X_c + u_c \quad [1]$$

Con ella buscamos establecer correlaciones entre los factores F y el índice γ en la comuna c . En este caso, los parámetros de interés están determinados por β_1 , β_2 , β_3 y β_4 . Finalmente, X representa un vector de controles (variables sociodemográficas) y u el error estándar.

⁶ Los coeficientes utilizados para la predicción a nivel de la unidad de análisis se muestran en la Tabla A4.

⁷ Esto requiere que las variables estén estandarizadas para que no exista influencia en la magnitud de las variables. Las variables predichas del análisis factorial ya se encuentran en términos de desviaciones estándar (Schwarz 2008; Mackles 2012; Visauta y Martori 2003).

3.

RESULTADOS Y ANÁLISIS

En esta sección presentamos los resultados principales. Iniciamos con algunos hechos estilizados sobre relaciones entre los índices generados, la victimización y el subreporte. Luego presentamos los resultados del análisis factorial con la subsecuente generación de los clústeres y su correspondiente caracterización. Finalmente, mostramos regresiones lineales que den indicios de la correlación entre los índices de percepción generados y las dimensiones predichas del análisis factorial.

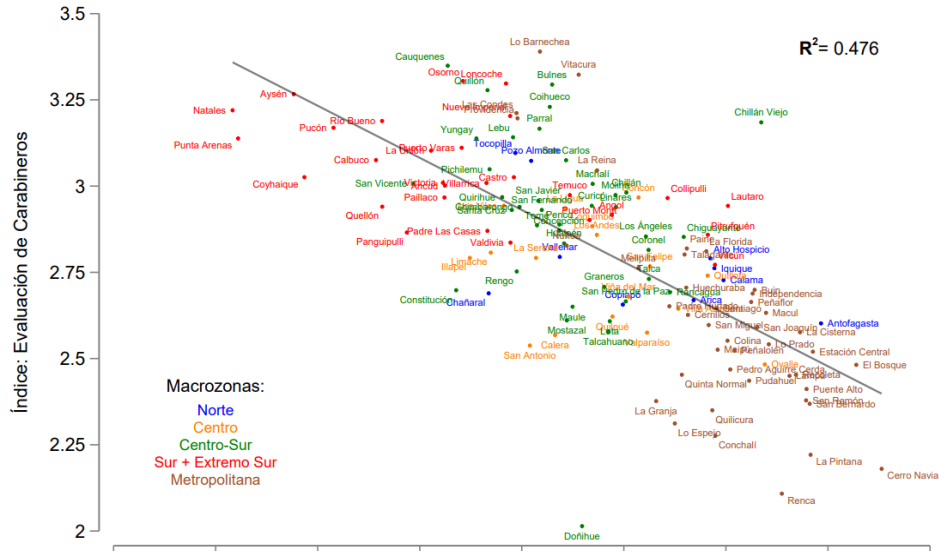
Hechos estilizados: índices, delincuencia y confianza

La preocupación por la delincuencia es un fenómeno latente y largamente estudiado en Chile. Estudios muestran que dicha preocupación y sensación de inseguridad pudiese modificar la estructura de preferencias y comportamientos (Felson et al. 2017; Gamarra et al. 2024). Esto se observa en la relación entre evaluación e inseguridad en las instituciones en la Figura 1.

La figura muestra la relación entre la evaluación de la labor de Carabineros y la sensación de inseguridad en lugares públicos. La correlación es elevada (R^2 de la regresión bivariada igual a 0.48, siendo la correlación de aproximadamente 0.7). Ello indica que el índice que mide la sensación de inseguridad es capaz de explicar el 50% de la varianza en la evaluación de la labor de Carabineros contra la delincuencia. Además de la importancia de dicha relación, patrones geográficos son visibles en la Figura 1. La Región Metropolitana es aquella que posee la peor evaluación de Carabineros, así como una alta sensación de inseguridad. Específicamente, en el sector poniente de Santiago, la evaluación de la labor de Carabineros se ubica en alrededor de 2 puntos en el índice (evaluación Mala). Por otro lado, excluyendo la Región Metropolitana, se observa un patrón de norte a sur, con las comunas de la macrozona norte en niveles similares a la región Metropolitana; luego siguen la macrozona centro, centro sur y sur, y el extremo sur.

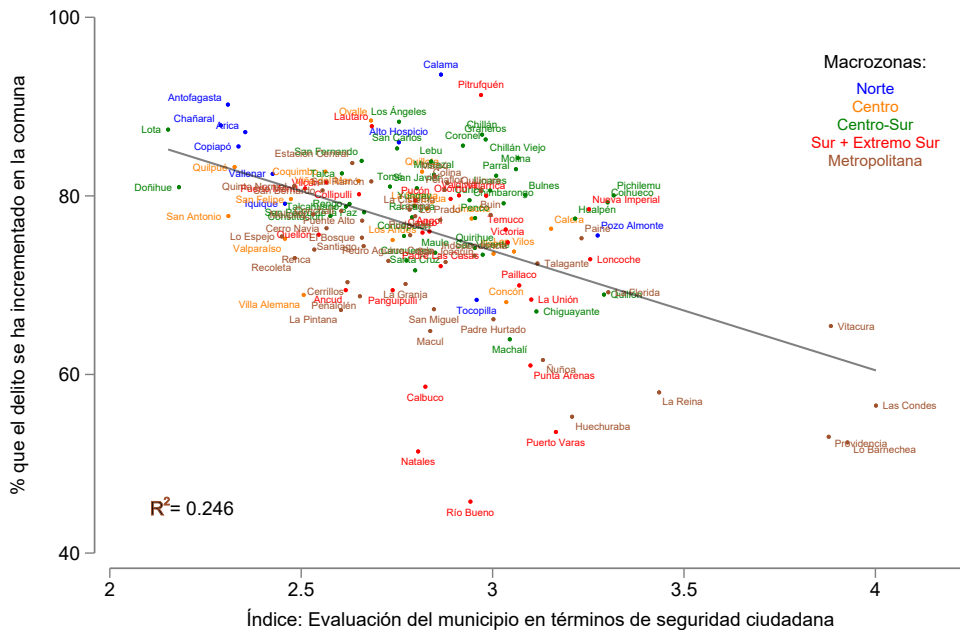
Esta evaluación no solamente se restringe a aquellas instituciones que tienen directa relación con la delincuencia; se extiende también al municipio, como muestra la Figura 2.

FIGURA 1. Correlación entre la evaluación de la labor de Carabineros y la inseguridad en lugares públicos, ENUSC 2023



Nota: La línea sólida negra representa la pendiente de la relación (es decir, el coeficiente β de la regresión incondicional).
 Fuente: Elaboración propia en base a datos ENUSC 2023.

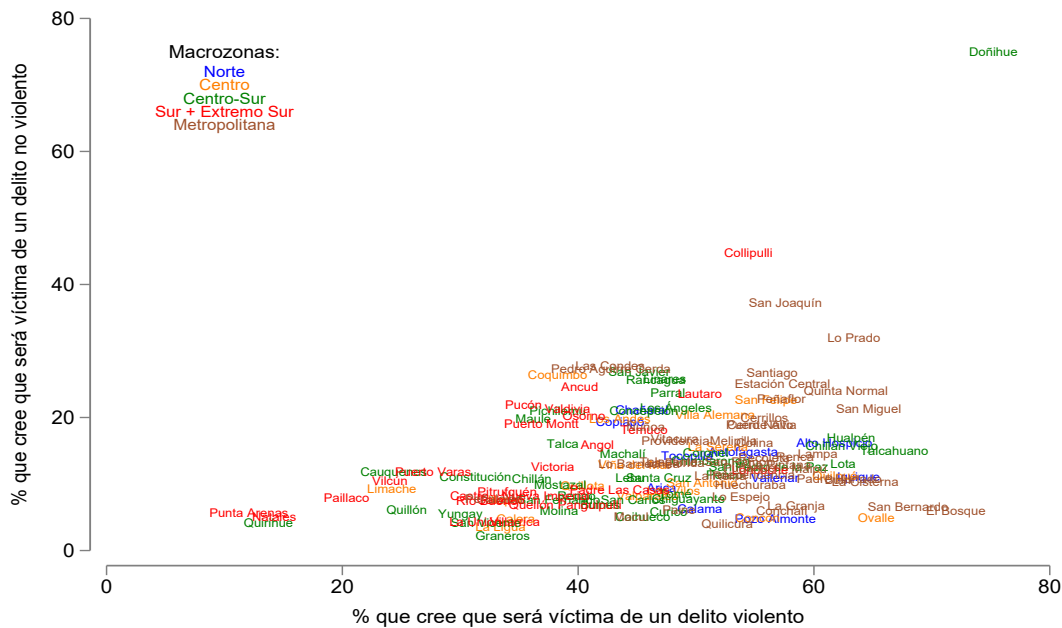
FIGURA 2. Correlación entre la percepción de aumento de la delincuencia en la comuna y la evaluación del municipio sobre su contribución a la seguridad ciudadana, ENUSC 2023



Nota: La línea sólida negra representa la pendiente de la relación (es decir, el coeficiente β de la regresión incondicional).
 Fuente: Elaboración propia en base a datos ENUSC 2023.

La Figura 2 muestra que las comunas que peor evalúan a su municipio —en términos de su labor para una mayor seguridad ciudadana— son aquellas que más fuertemente perciben que la delincuencia se ha incrementado en su comuna. La relación —aunque menos prominente que la Figura— muestra una correlación elevada. El índice de evaluación del municipio es capaz de explicar el 24% de la variación en la percepción del aumento de la delincuencia en la comuna. Nuevamente, un patrón geográfico se aprecia en la Figura 2. Las comunas con mejor evaluación del municipio (comunas de altos ingresos de la Región Metropolitana) poseen una menor percepción de peligro. Esto contrasta fuertemente con las comunas en la macrozona norte, donde una alta percepción de peligro —más del 80% para casi todas las comunas— se correlaciona con una baja evaluación del municipio en términos de seguridad ciudadana. Esta elevada percepción de peligro incide en las expectativas, además de alterar los comportamientos y preferencias (Gamarra et al. 2024). La Figura 3 indaga en esto para delitos violentos y no violentos.

FIGURA 3. Scatterplot entre el porcentaje de encuestados que cree que será víctima de un delito violento y no violento, ENUSC 2023



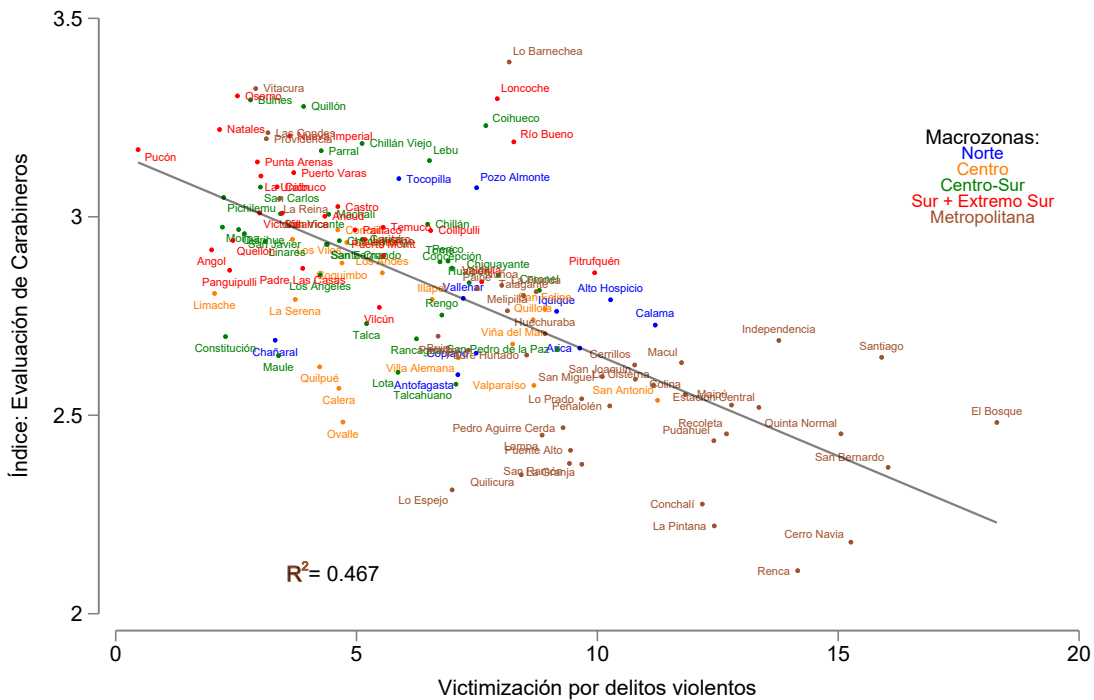
Nota: La línea sólida negra representa la pendiente de la relación (es decir, el coeficiente β de la regresión incondicional).
Fuente: Elaboración propia en base a datos ENUSC 2023.

La Figura 3 muestra que, a pesar de que la victimización sea más baja para delitos violentos, los encuestados creen probable ser víctimas de un delito violento en los próximos 12 meses. Para la mayor parte de las comunas, este porcentaje supera el 40% (eje horizontal). Nuevamente, en las co-

munas de la Región Metropolitana es más probable tener la expectativa de victimización por delitos violentos.⁸

A grandes rasgos, las correlaciones sugieren que las percepciones de los ciudadanos están conectadas. Es decir, lugares con mayor percepción de inseguridad son aquellos en los que, de la misma forma, se confía menos o se evalúa peor a la policía. De cualquier modo, que exista una relación entre las percepciones, no significa que ellas estén directamente conectadas con los niveles de violencia al que efectivamente están expuestos los hogares. En tal sentido, diversos estudios indican que la percepción de la delincuencia (o, específicamente, del incremento de la delincuencia) no está asociada de manera directa con el nivel de delincuencia efectivo (Buil-Gil et al. 2021; Hayward 2007; Navarrete-Hernández et al. 2023). La percepción sobre el fenómeno de la delincuencia es heterogénea y no se restringe a un solo indicador ni distintos tipos de delitos impactan del mismo modo las preferencias, percepciones y pronósticos de los hogares.

FIGURA 4. Correlación entre el porcentaje de hogares victimizados y la evaluación de la labor de Carabineros, ENUSC 2023



Nota: La línea sólida negra representa la pendiente de la relación (es decir, el coeficiente β de la regresión incondicional).
Fuente: Elaboración propia en base a datos ENUSC 2023.

⁸ El caso más sorprendente es el Doñihue, que presenta una relación casi perfecta y bastante elevada, con aproximadamente 75% de sus encuestados que pronostican ser víctimas de delitos violentos y no violentos en los próximos 12 meses.

La Figura 4 ilustra un caso en particular: en comunas en las que la probabilidad de ser victimizado por algún delito violento es mayor, la evaluación de la labor de Carabineros es notoriamente baja. La variación de la probabilidad de ser víctima de delitos violentos captura casi el 50% de la variabilidad en la evaluación de Carabineros a través de las comunas. Este es solo un aspecto, pues la concentración de determinados delitos en comunas particulares puede generar dinámicas entre esas comunas que podrían ser ignoradas al utilizar datos desagregados (i.e., individuales) que no toman en cuenta la anidación de los eventos relacionados con la delincuencia (Riascos Villegas et al. 2023; Brantingham et al 2017; Bruinsma y Johnson 2018; Groff 2018).

En este contexto la comuna adquiere relevancia en tanto existen relaciones que se pueden representar dicho nivel de agregación. Por ejemplo, cómo en determinadas comunas que poseen ciertas características, se concentran determinados tipos de delitos, y que eso conlleve a que —para esa comuna, en una evaluación promedio de las respuestas de los hogares— la percepción del rol y labor del municipio o Carabineros sea deficiente. Más aún, a través de las figuras presentadas se ha observado que existe una relación a un nivel incluso más agregado: conexiones regionales, cuyo caso más saliente es el de la Región Metropolitana.

Así, con el propósito de desentrañar dichas constelaciones espaciales, empleamos técnicas exploratorias que, sin hipótesis previa, permiten indagar en las asociaciones y dinámicas entre comunas y delitos. Observamos la distribución heterogénea del delito: grupos de comunas que están más expuestos a determinados eventos (delitos). Como detallaremos más adelante, esta agrupación no solo se relaciona con aspectos de victimización, sino que esos grupos de comunas también poseen percepciones, comportamientos y evaluaciones similares entre sí.

Análisis factorial y *clustering*

Para la construcción del número óptimo de clústeres se utilizaron criterios estándar de la literatura (Mackles 2012). En la sección sobre el método hemos presentado los estadígrafos utilizados para la selección del número óptimo, y el Anexo 1 presenta con mayor detención el proceso y los estadígrafos que nos permiten concluir que 6 es el número óptimo de clústeres (Figura A1). Las comunas en cada agrupación se pueden apreciar en la Tabla 1, y la representación de cada macrozona en los clústeres es visible en la Figura A2.

TABLA 1. Comunas en cada clúster

Clúster	Comunas
C1 (14,18%)	Ancud - Calera - Chañaral - Chillán - Collipulli - Concepción - Curicó - Lautaro - Paillaco - Penco - Pozo Almonte - Puerto Montt - Quillota - Rancagua - Río Bueno - San Antonio - San Felipe - Talcahuano - Tomé
C2 (11,94%)	Cerrillos - Coihueco - Estación Central - Graneros - Osorno - Padre Las Casas - Paine - Panguipulli - Pedro Aguirre Cerda - Quillón - San Carlos - San Fernando - San Joaquín - San Ramón - Victoria - Vilcún
C3 (17,91%)	Alto Hospicio - Arica - Cerro Navia - Colina - Conchalí - Independencia - Iquique - La Cisterna - La Florida - La Granja - La Pintana - Lampa - Lo Prado - Macul - Maipú - Melipilla - Pudahuel - Puente Alto - Quilicura - Quinta Normal - Recoleta - Renca - San Bernardo - Valparaíso
C4 (12,69%)	Castro - Chiguayante - Coronel - El Bosque - Lebu - Loncoche - Machalí - Ovalle - Peñaflor - Pitrufquén - Puerto Varas - Rengo - San Pedro de la Paz - Santa Cruz - Temuco - Valdivia - Viña del Mar
C5 (11,19%)	Calama - Concón - Huechuraba - La Reina - Las Condes - Lo Barnechea - Los Ángeles - Padre Hurtado - Peñalolén - Providencia - Quellón - San Miguel - Santiago - Vitacura - Ñuñoa
C6 (32,09%)	Angol - Antofagasta - Buin - Bulnes - Calbuco - Cauquenes - Chillán Viejo - Chimbarongo - Constitución - Copiapó - Coquimbo - Doñihue - Hualpén - Illapel - La Ligua - La Serena - La Unión - Limache - Linares - Lo Espejo - Los Andes - Los Vilos - Lota - Maule - Molina - Mostazal - Natales - Nueva Imperial - Parral - Pichilemu - Pucón - Punta Arenas - Quilpué - Quirihue - San Javier - San Vicente - Talagante - Talca - Tocopilla - Vallenar - Villa Alemana - Villarrica - Yungay

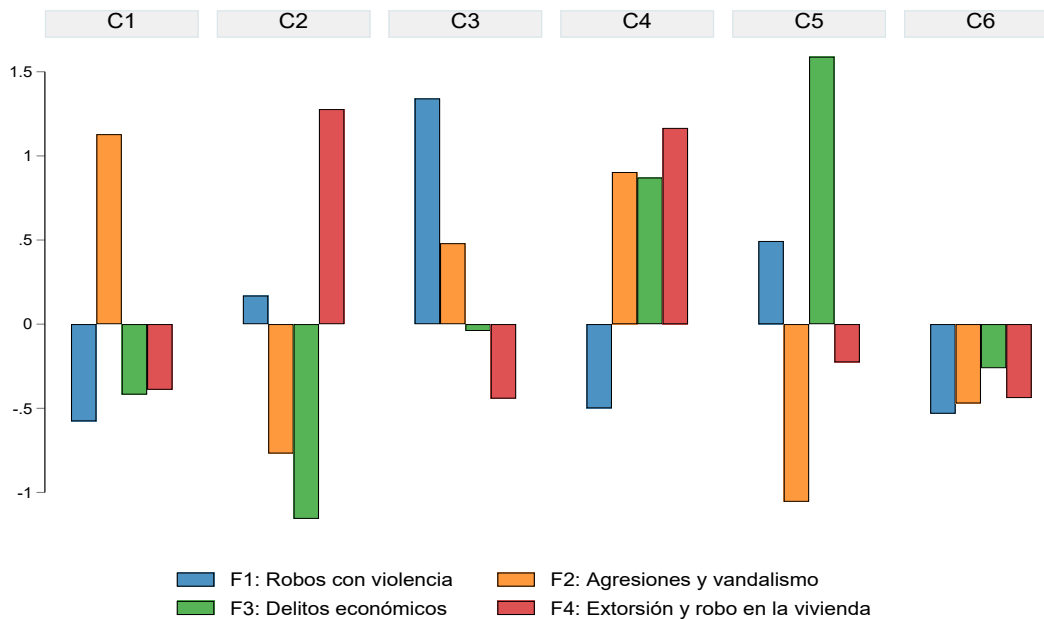
Nota: Las variables utilizadas en la generación de los clústeres fueron los 4 primeros factores del análisis factorial por componente principal —con rotación ortogonal varimax— utilizando la proporción de hogares victimizados a nivel comunal para 16 categorías de delito consultadas en ENUSC 2023. El método de clusterización fue el de k-medias. Para mayor detalle en cómo se generó la clusterización, revisar Anexos 1. Se muestra también el porcentaje de comunas que contiene cada clúster.

La Figura 5 y la Tabla 2 dan indicios sobre la relación entre los factores y los clústeres, entre la concentración y preponderancia de ciertos delitos dentro de cada clúster de comunas. En esta se observa que C1 —predominantemente con comunas de la zona centro sur y extremo sur— posee niveles elevados del factor *agresiones y vandalismo* (aproximadamente 1 desviación estándar por encima). Para los otros factores dicho clúster se encuentra bajo el promedio. En C2 la incidencia del factor 4 —con varias comunas de la Región Metropolitana y la zona centro sur y extremo sur y que guarda relación con los robos de vivienda y delitos por extorsión— es alta. En este mismo clúster, los factores 2 y 3 —*delitos económicos* y *agresiones y vandalismo*— están considerablemente bajo el promedio, mientras que el factor *robo con violencia* está ligeramente sobre el promedio. C3, compuesto principalmente por comunas de la Región Metropolitana, posee una fuerte incidencia del factor *robo con violencia* (superior a 1 desviación estándar sobre el promedio). La incidencia del factor 2, *agresiones y vandalismo*, es menos marcada pero también positiva (alrededor de 0.4 desviación estándar). El resto de los factores se encuentra bajo el promedio. C4 —principalmente con comunas de la zona centro sur de Chile— posee una alta incidencia de los factores *agresiones y vandalismo*, *delitos económicos*, y *extorsión y robo de vivienda*. Estos factores son superiores a 0.8 desviaciones estándar sobre el promedio. C5 —predominantemente con comunas de la Región Metropolitana— muestra incidencia del factor *delitos económicos* (casi 1.5 desviaciones estándar). Casualmente, en dicho clúster se encuentran las

comunas con mayores ingresos de la Región Metropolitana. Este clúster también posee niveles del factor *robo con violencia* de alrededor de 0.5 desviaciones estándar sobre el promedio, y un nivel de *agresiones y vandalismo* sustancialmente bajo el promedio (aproximadamente 1 desviación estándar). Por último, C6 —con mayor cantidad de comunas de la zona centro sur— posee niveles inferiores al promedio en los 4 factores analizados.

En síntesis, el clúster 1 posee niveles de *agresiones y vandalismo* sobre el promedio. El clúster 2 posee elevados niveles de *extorsión y robo a la vivienda*, aunque *delitos económicos* y *agresiones y vandalismo* se ubican bajo el promedio. El clúster 3 es el que mayor victimización posee por los delitos conocidos como de ‘alta connotación social’, bajo la dimensión de *robos con violencia*, en nuestro caso. El clúster 4 posee niveles considerables de *agresiones y vandalismo*, *delitos económicos*, y *extorsión y robo a la vivienda*, aunque con *robo con violencia* bajo el promedio. El clúster 5 presenta elevados niveles de *delitos económicos* y sobre el promedio en *robo con violencia*. Finalmente, el clúster 6 está bajo el promedio en los 4 factores.

FIGURA 5. Promedio de los factores para cada clúster



Nota: Los factores fueron generados para cada unidad de análisis (comuna) de acuerdo con el método *regression scoring*. La Tabla A3 presenta los coeficientes utilizados para la generación de los clústeres. Los factores están estandarizados y no correlacionados entre ellos.

TABLA 2. Distribución por macrozonas y dimensiones del delito

	C1	C2	C3	C4	C5	C6	Promedio total
Panel A: Distribución por macrozonas (%)							
Norte	10,5	0,0	12,5	0,0	6,7	9,3	7,5
Centro	21,1	0,0	4,2	11,8	6,7	20,9	12,7
Centro-Sur	36,8	31,3	0,0	41,2	6,7	44,2	29,1
Sur + Extremo Sur	31,6	31,2	0,0	35,3	6,7	18,6	19,4
Metropolitana	0,0	37,5	83,3	11,8	73,3	7,0	31,3
Panel B: Dimensiones del delito							
Robos con violencia	-0,58+	0,17	1,34*	-0,50+	0,49	-0,53+	0,00
Agresiones y vandalismo	1,13*	-0,77*	0,48	0,90*	-1,06*	-0,47	0,00
Delitos económicos	-0,42	-1,16*	-0,04	0,87*	1,59*	-0,26	0,00
Extorsión y robo a la vivienda	-0,39	1,28*	-0,44	1,17*	-0,23	-0,44	0,00
Número de comunas	19	16	24	17	15	43	134

Nota: El Panel A ilustra la proporción de comunas que pertenecen a las seis macrozonas, para cada clúster. Esta información puede ser leída con respecto a la columna que indica la proporción promedio: si una macrozona representa un mayor porcentaje en un determinado clúster con respecto a su proporción promedio en el total de observaciones, esa macrozona está sobrerrepresentada en dicho clúster. Por ejemplo, de la columna que refiere al clúster 1, se puede afirmar que la macrozona Norte está sobrerrepresentada. El Panel B muestra el promedio de cada dimensión proveniente del análisis factorial, para cada clúster. Se resaltan los valores que superan 0,5 desviaciones estándar, siguiendo una interpretación tradicional de tamaño de efecto (d de Cohen). En ese sentido, valores superiores a 0,5 pueden ser considerados como un efecto mediano, mientras que aquellos superiores a 0,7 como efectos notables: + superior a 0,5; * superior a 0,7. En ese sentido, se puede interpretar que las 19 comunas que pertenecen al clúster 1 poseen, en promedio, elevados niveles de agresiones y vandalismo, pero menos robos con violencia.

Existe entonces una relación entre tipos de delito (asociados dentro de los factores) cuya distribución es heterogénea en diversas comunas (o agrupaciones de comunas, posibles de dilucidar a través de los clústeres). La Tabla 3 presenta una caracterización de los clústeres en términos de los índices generados. El clúster 3, notoriamente el de mayor victimización violenta, es la agrupación con menor confianza en las instituciones (Carabineros, PDI y Fiscalía y Ministerio Público), mayor pronóstico de ser víctima por delitos violentos, mayor percepción de aumento de la delincuencia en el barrio y peor evaluación de la labor de Carabineros y el municipio en términos de la provisión de seguridad ciudadana. A su vez, también es la agrupación donde se siente inseguridad en la mayor cantidad de lugares públicos, donde más incivildades y desórdenes se observan en los barrios, y donde con mayor prominencia los encuestados afirman que su opinión sobre la delincuencia en la comuna está basada en su experiencia personal.

En los demás clústeres la relación es menos prominente pero no por ello de menor interés. En algunos casos, las relaciones pueden incluso llegar a pensarse como contraintuitivas. Por ejemplo, el clúster 5 posee niveles de victimización por delitos económicos notablemente sobre el promedio y ligeramente por encima del promedio para el caso de robo con violencia. A su vez, es el clúster en el que se encuentran algunas de las comunas con mayor nivel socioeconómico. Sin embargo, en este clúster, la confianza en Carabineros y la PDI es la más alta, pero exhibe niveles elevados sobre la expectativa de ser

víctimas de delitos violentos y no violentos. Es decir, a pesar de la alta confianza y buena evaluación de las instituciones que poseen, este clúster está compuesto por las comunas que pronostican con mayor probabilidad ser víctimas de delitos.

Por otra parte, el clúster 6 —compuesto por casi un tercio de las comunas analizadas— es el clúster con menores niveles de victimización para todos los delitos analizados. Notablemente, es también el clúster que menor percepción de aumento de la delincuencia en el barrio posee, donde menor inseguridad en lugares públicos se percibe, y donde hay menor expectativa de los hogares de ser víctimas de delitos violentos. Asimismo, es el clúster en el que menos incivildades y desórdenes se aprecian.

TABLA 3. Caracterización de los clústeres utilizando los índices generados

Índices	C1	C2	C3	C4	C5	C6
Percepción de aumento de la delincuencia (%)						
País	0,90	0,90	0,88	0,89	0,86	0,90
Comuna	0,77	0,77	0,77	0,77	0,67	0,77
Barrio	0,47	0,51	0,58	0,49	0,51	0,45
Pronóstico de ser víctima (%):						
Delito violento	0,45	0,43	0,57	0,47	0,50	0,42
Delito no violento	0,16	0,14	0,13	0,12	0,15	0,14
Confianza en las instituciones (%):						
Carabineros	0,61	0,59	0,55	0,62	0,68	0,62
PDI	0,74	0,69	0,69	0,73	0,78	0,72
Fiscalía y Ministerio Público	0,38	0,32	0,28	0,35	0,32	0,37
Evaluación de la labor de instituciones:						
Carabineros (1 a 5)	2,85	2,84	2,50	2,86	2,91	2,89
Municipio (1 a 5)	2,77	2,87	2,70	2,88	3,19	2,82
Fuente de información sobre la delincuencia en la comuna:						
Experiencia personal (%)	0,07	0,10	0,17	0,09	0,13	0,07
Otros índices:						
Inseguridad en lugares públicos (0 a 16)	4,49	4,86	6,17	4,86	4,91	4,21
Medidas adoptadas (0 a 9)	0,75	0,71	0,74	0,82	0,60	0,66
Apoyo entre vecinos (1 a 4)	2,45	2,49	2,46	2,44	2,49	2,46
Incivildades (0 a 7)	0,48	0,66	1,11	0,57	0,57	0,44
Desórdenes (0 a 8)	1,30	1,54	2,47	1,40	1,57	1,15

Nota: Se presentan los promedios de cada indicador, para cada clúster. En el caso de las variables dicotómicas, el promedio es directamente la proporción. Para la percepción de aumento de la delincuencia, se considera los hogares que respondieron que la delincuencia se ha incrementado, en país, comuna y barrio. Para pronóstico de ser víctima de delitos, se considera los hogares que afirman poder ser víctima de alguna categoría de delito, diferenciando de si esta es violento o no violenta. Confianza en las instituciones representa la proporción de hogares que afirma tener Mucha o Bastante Confianza, de la muestra que afirma conocer las labores en términos de seguridad ciudadana, para cada institución. También se ilustra el porcentaje de hogares que anuncia que la fuente de información sobre la delincuencia en su comuna es su experiencia personal. Finalmente, los demás índices ilustran el promedio de los indicadores generados, así esclarecidos en la sección de Datos. Mientras más alto el índice en evaluación de la labor de instituciones, mejor evaluada la labor de dicha institución. Para las variables dentro de otros índices, mientras más grande el valor, significa mayor inseguridad en lugares públicos, mayor cantidad de medidas adoptadas por los hogares debido a la delincuencia, mayor apoyo entre vecinos, mayor presencia de incivildades y desórdenes en su barrio.

Los efectos de las dimensiones del delito

El rol de las percepciones sobre la delincuencia y sus causas es un fenómeno estudiado en Chile (e.g., Cadena-Urzuía et al. 2023; Luneke 2016), aunque su relación con la victimización es subanalizada. El estándar es que existe una distribución espacial heterogénea del delito: ciertos tipos de delitos se concentran en determinadas comunas con características particulares (Weinborn et al. 2017; Gamarra et al. 2024). Se ha esgrimido que el nivel de análisis comunal posibilita identificar efectos que no solamente serían ignorados con otra unidad de análisis, sino que dicha agregación es significativa (no solamente por la existencia de esta distribución heterogénea, sino por las características compartidas dentro de cada comuna que poseen los hogares con respecto a sus experiencias y percepciones sobre la delincuencia).

En este contexto, utilizando los índices generados previamente y los resultados del análisis factorial es posible aproximarse a la relación entre las percepciones (sobre el incremento de la delincuencia, el pronóstico de ser víctima de delitos, la confianza y evaluación de las instituciones de seguridad) y la facticidad de la delincuencia (identificada por las declaraciones de las personas de haber sido víctimas de delitos) a través de métodos que permitan la interrelación entre diversas variables.

Para este propósito, empleamos los factores obtenidos como posibles predictores (variables independientes) de un conjunto de variables que ilustren la percepción de los hogares en relación con fenómenos relacionados a la delincuencia, así como con variables que indiquen aspectos más factuales, como el subreporte y la revictimización (variables dependientes). Los resultados son presentados en la Tabla 4. Los coeficientes provienen de una regresión OLS para discernir la correlación entre los factores y un conjunto de 11 variables, a nivel comunal, que incluyen los factores (dimensiones del delito) y algunos controles sociodemográficos.⁹ Para las variables (factores), el incremento marginal se encuentra en términos de desviaciones estándar.

El factor I (*robo con violencia*) posee un efecto estadísticamente significativo de casi 3 puntos porcentuales en la probabilidad de que las comunas no reporten a las autoridades correspondientes el haber sufrido algún delito violento. A su vez, este factor también posee un impacto significativo positivo sobre la posibilidad de que un hogar sea revictimizado por un delito no violento. Es decir, comunas con mayor incidencia de los delitos contenidos en el factor I poseen una mayor probabilidad de no reportar los delitos violentos sufridos, y de ser víctimas más de una vez de delitos no violentos. Esto, de forma esperable, se asocia con que este mismo factor posee un efecto sustancial sobre la probabilidad de que más hogares en una misma comuna crean sufrir al menos un delito violento en los próximos 12 meses, o perciban más delincuencia en su barrio y confíen menos en Carabineros y el Ministerio Público.

⁹ No se recurre a la interpretación de dichos parámetros debido a que dichas variables (i.e., % de hogares de bajo nivel socioeconómico) no se relacionan perfectamente con la distribución sociodemográfica en esa comuna. Sin embargo, creemos que es útil controlar por las características de los hogares que efectivamente responden la ENUSC 2023. Por ello, se consideran para la estimación, pero sus coeficientes no son mostrados.

El factor 2 (*agresiones y vandalismo*) posee un impacto significativo en la probabilidad de no reportar un delito no violento, ser revictimizado por algún delito (ya sea violento o no violento) e incrementa la probabilidad de que los hogares en una misma comuna indiquen ser víctimas probables de delito violento en los próximos 12 meses.

Con menores relaciones significativas que los demás factores, por una parte, el factor 3 posee un efecto significativo y positivo sobre la revictimización por delito violento y el pronóstico de ser víctima de un delito violento. Por su parte, el factor 4, posee un impacto positivo en la revictimización por delito violento y aumenta la percepción de incremento de la delincuencia en el barrio. Este último factor contiene los delitos asociados a robos en la vivienda.

Los factores y el conjunto de controles son capaces de capturar gran parte de la varianza de las variables dependientes a través de las comunas (en algunos casos sobrepasando el 50%). Esto, no obstante, no significa que exista un impacto causal de los factores sobre las variables anunciadas, debido a que —si bien discernir los efectos entre los factores es un buen ejercicio porque no correlacionan entre sí— no existen razones claras para suponer que dichos factores no están correlacionados con variables no observadas que no están siendo incluidas en el modelo, lo que posiblemente sesga los coeficientes y limita su interpretación en términos causales.

TABLA 4. Relación lineal entre los factores generados e índices sobre subreporte, victimización, pronóstico y percepciones sobre la delincuencia y confianza en instituciones

	Subreporte %		Revictimización %		Pronóstico de delito	Percepción de incremento de la delincuencia		Confianza y evaluación de instituciones		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Delito violento	Delito no violento	Delito violento	Delito no violento	Delito violento	Comuna	Barrio	Carabineros	Fiscalía y Ministerio Público	Municipio
Factor 1	2,903*	1,301	-1,134	2,341**	6,797***	0,0910	6,191***	-2,750***	-4,599***	-0,0243
	(1,66)	(1,36)	(-0,87)	(2,49)	(5,70)	(0,10)	(6,04)	(-3,46)	(-4,75)	(-0,84)
Factor 2	0,170	2,005***	1,959*	3,800***	2,052**	0,520	0,688	-0,245	-0,793	0,0578***
	(0,12)	(2,66)	(1,96)	(5,28)	(2,25)	(0,72)	(0,88)	(-0,40)	(-1,07)	(-2,62)
Factor 3	1,880	0,178	0,115	1,792**	1,822*	-0,677	-1,100	0,352	0,115	0,0379
	(1,18)	(0,20)	(0,10)	(2,08)	(1,67)	(-0,78)	(-1,17)	(0,48)	(0,13)	(1,44)
Factor 4	1,082	-0,123	2,156**	0,455	0,731	0,958	1,494**	0,0111	-0,0820	0,00617
	(0,87)	(-0,18)	(2,32)	(0,68)	(0,86)	(1,41)	(2,04)	(0,02)	(-0,12)	(0,30)
N	129	130	134	134	134	134	134	134	134	134
R ²	0,091	0,121	0,116	0,353	0,421	0,275	0,515	0,534	0,304	0,515
Test F	1,064	1,476	1,459	6,046	8,058	4,205	11,79	12,70	4,840	11,79

Nota: La tabla ilustra los coeficientes de una regresión lineal, sobre cada variable dependiente ilustrada en las columnas, utilizando como predictores los factores generados del análisis factorial y añadiendo como variables de control (coeficientes no mostrados): % de hogares de alto NSE, % de hogares de bajo NSE, % de hogares pertenecientes a pueblos indígenas, % de mujeres y % de personas entre 40 y 59 y 65 a más años, y % de migrantes. La unidad de análisis es la comuna, por ende, la cantidad de comunas en cada estimación refiere a las observaciones de dicha estimación. Todas las variables dependientes son continuas. Las variables representan una proporción [0, 100], excepto la columna (10) cuyo mínimo es 1 y máximo es 5 (promedio de la evaluación de Carabineros a nivel comunal, en base a ítems por hogar). Para mayor detalle en la generación de las variables dependientes, revisar la sección de Datos. El estadístico t se presenta entre paréntesis; * p < 0.1, ** p < 0.05, *** p < 0.01.

4.

LÍNEAS DE EXPLORACIÓN Y PROFUNDIZACIÓN

Puesto que este artículo tiene una finalidad exploratoria, en esta sección identificamos líneas de investigación futura y de análisis de política pública que derivan de los resultados producidos por los métodos aplicados.

- En buena parte de los resultados, la Región Metropolitana destaca como una zona geográfica con alta variabilidad y valores extremos. En este caso, es preciso investigar en profundidad las comunas de la zona poniente de la ciudad de Santiago —incluso en lugares específicos de las comunas. Estas comunas presentan alta inseguridad y baja evaluación de la labor de Carabineros, que esperan ser víctimas de delitos violentos, y que muestran diferencias relevantes por nivel socioeconómico de las comunas. La exploración de estas diferencias comunales en la Región Metropolitana debe ser parte de estudios detallados pero integrativos que se interroguen, por ejemplo, por la movilidad criminal en la ciudad, por sus *peaks* estacionales, por la previsible presencia de crimen organizado en ellas.
- Es preciso explorar en profundidad también en las expectativas de las personas en relación con las labores de seguridad del municipio. La tarea general de seguridad pública pertenece al Estado, sin embargo, las personas esperan más de sus municipios en materias de seguridad. Más allá de la seguridad privada que los municipios puedan (o no) tener, estos pueden desplegar distintas formas de prevención, desarrollo de espacios públicos (plazas, iluminación, remodelación) e inclusión social que limiten la comisión de delitos. Es necesario explorar las expectativas ciudadanas a nivel comunal a este respecto y relacionarlas con las posibilidades de implementación municipal en cada caso.
- Los clústeres de comunas construidos por medio de los procedimientos metodológicos descritos agrupan comunas de distintos ámbitos geográficos. Al relacionar esos clústeres con los factores, se aprecian algunos patrones geográficos, pero también una alta variabilidad. Es preciso explorar estas agrupaciones con mayor profundidad en sus patrones visibles, pero especialmente en patrones ocultos que pueden estar tras la relación entre comunas reunidas en un clúster, pero de distribución espacial variable. Justamente en estos patrones ocultos se pueden identificar explicaciones diversas y hasta ahora desconocidas de las dinámicas delincuenciales a nivel territorial y de sus modos de contención.
- Este análisis de clústeres también exhibe (clúster 3) una relación negativa entre la experiencia de la delincuencia a nivel comunal y la confianza institucional. En las últimas décadas (salvo por una caída en el estallido social), las instituciones de orden público han presentado una alta confianza

institucional a nivel nacional. El análisis de clústeres muestra que esta confianza no es generalizada. El clúster 3 tiene una alta representación de la Región Metropolitana (83,3%) y el factor más asociado a él es el de robos con violencia (1,34). La confianza en las instituciones de orden público está entonces asociada principalmente a los robos con violencia. Esta es una asociación que es preciso investigar en profundidad para introducir variación en la idea de que las instituciones policiales gozan de confianza ciudadana a todo evento.

- Asimismo, el clúster 6 agrupa las comunas en las que todos los factores de delito son bajos. La investigación sobre el crimen se centra de manera regular y obvia en las prácticas criminales, por tanto, en los espacios donde estas prácticas tienen lugar. La pregunta que abre esta agrupación comunal con índices de delincuencia bajo el promedio es qué tienen de común estas comunas que produce estos índices moderados de delincuencia. Esta pregunta abre un campo de investigación amplio, desde buenas prácticas ciudadanas, municipales, policiales, hasta cuestiones de composición demográfica, patrones de migración interna, diseños urbanos, políticas de inclusión. La identificación de los elementos compartidos entre estos espacios territoriales puede sin duda ayudar al diseño de políticas de distinto tipo para la reducción del crimen en los ámbitos indicados u otros. La investigación sobre las formas de delito y su control debe entonces aprender no solo de los lugares donde el delito se comete, sino también de los espacios donde los índices hacen visible una contención.
- Un área de investigación relevante de formato más cualitativo se relaciona con la asociación entre ser víctima de algún tipo de delito y el subreporte del delito a las instituciones policiales. En una investigación anterior (Gamarra et al. 2024) hemos observado que incluso el delito violento muestra una alta probabilidad de subreporte. Los resultados de los ejercicios desarrollados en este artículo corroboran esta tendencia. La percepción de que la policía es inefectiva en la resolución de los casos disminuye paulatinamente la probabilidad de reportar un nuevo crimen que se ha sufrido, incluso si este es de tipo violento. Considerando esto, es preciso indagar en la sociología del subreporte, no solo en sus motivaciones y razones explícitas de tipo individual, sino también en la interrelación de factores territoriales, interpersonales, institucionales y normativos que anclan esta conducta en determinados grupos de personas.

Las líneas de investigación posibles no se agotan en las formulaciones descritas. Sin embargo, ellas derivan explícitamente del ejercicio metodológico realizado. Su indagación futura puede permitir nuevos conocimientos en las dinámicas del crimen en Chile y, con ello, políticas públicas más eficientes en el control y prevención del crimen.

5.

CONCLUSIONES

En este artículo hemos realizado una investigación exploratoria inicial del delito a nivel comunal en Chile a partir de los datos de ENUSC 2023, encuesta que posee información de aproximadamente 50.000 hogares en 136 comunas de las 346 del país. Realizada mediante técnicas estadísticas avanzadas, la exploración consistió en determinar si existe alguna correlación entre la ocurrencia de delitos violentos y no violentos a nivel espacial comunal y algunas variables centrales de interés como percepción de inseguridad, la evaluación de la actuación de la policía, la probabilidad de ser víctima de un delito violento y determinadas variables construidas por análisis factorial denominadas robo con violencia, agresión y vandalismo, delitos económicos, extorsión y robo a la vivienda. De esto, hemos sugerido líneas de profundización y exploración. Algunas conclusiones derivadas del análisis son las siguientes.

La percepción de inseguridad continúa siendo un factor relevante en la constelación actual del crimen. Independiente de la relación entre esta percepción y el aumento real de delitos (regularmente hay una distancia entre estas cifras), la percepción de inseguridad influye negativamente en la evaluación en Carabineros. Esto es especialmente crítico en el sector poniente de la Región Metropolitana, donde se concentra la mayor parte de los delitos violentos y donde la sensación de inseguridad está más concentrada. Algo similar acontece en algunas comunas de la macrozona norte como Antofagasta, Calama, Arica, Alto Hospicio y Copiapó. La mejor evaluación de Carabineros se ubica en la zona sur y austral de Chile. Asimismo, la percepción de inseguridad hace que municipios de las comunas de más altos ingresos del país (Lo Barnechea, Vitacura, Las Condes, Providencia) tengan una mejor evaluación por parte de la ciudadanía. Esto contrasta con las comunas de la zona norte, en las que el delito es más violento y la evaluación del municipio en términos de seguridad ciudadana es marcadamente baja.

La inseguridad, por tanto, no tiene una distribución homogénea a lo largo de las comunas del país. Ella está desigualmente distribuida. Considerando los datos para 136 de las 346 comunas de Chile, esta inseguridad se concentra en la zona poniente de la Región Metropolitana, en comunas que albergan centros urbanos relevantes de la zona norte, en comunas de la zona sur que han sido afectadas por movimientos étnonacionales y otras comunas específicas de la zona centro sur.

La desigualdad en la distribución de la inseguridad siempre tiene un componente asociado al estatus socioeconómico (como el que se expresa en la diferencia entre la zona poniente y oriente de la Región Metropolitana), sin embargo, esta no es la única variable. En la exploración realizada en esta

investigación, las comunas y zonas de mayor inseguridad parecen estar regularmente asociadas a una presión sistémica derivada de la producción de eventos críticos. La zona poniente de Santiago muestra mayores niveles de delitos asociados al crecimiento rápido del crimen organizado en los últimos años; comunas de la zona norte han sido especialmente afectadas por la migración masiva que tiene lugar desde 2019; las de la zona sur enfrentan, al menos desde hace una década, eventos permanentes de violencia étnica; y la zona centro sur ha sido fuertemente impactada por megaincendios forestales en 2017 y 2023. Estos eventos críticos no solo afectan a las personas directamente, sino que también introducen una presión que desborda la capacidad de las instituciones encargadas de la seguridad y de las instituciones sociales responsables de la inclusión en educación, salud y trabajo especialmente. Los eventos críticos, por tanto, parecen jugar un rol en la inseguridad asociada al crimen; son un factor que aumenta la inseguridad en las zonas específicas donde esos eventos críticos tienen su epicentro. Esto es algo que debe ser investigado más a fondo en sus consecuencias inmediatas y de mediano plazo.

La percepción de inseguridad actual en varios sentidos define también la percepción de inseguridad futura. Comunas como El Bosque, San Bernardo, Talcahuano y Ovalle asumen que serán víctimas de un delito violento en el futuro. También lo hacen otras como San Miguel, Quinta Normal, Lo Prado, San Joaquín, Collipulli y, especialmente, Doñihue, pero en estas, además de esperar ser víctimas de un delito violento, incluyen la expectativa de sufrir delitos no violentos en el futuro próximo. Esta doble expectativa no solo cambia las conductas presentes, sino que anticipa la sensación de riesgo y peligro a la que la población estará afectada en lo sucesivo, y con ello, adapta las conductas futuras a esos riesgos y peligros imaginados. Esto puede explicar el hecho que, aun cuando el delito pueda ser controlado y sus tasas reducidas en términos reales, la percepción de inseguridad actual se refuerce como memoria del pasado; que cada evento delictivo, aunque sea menor, active esa percepción y la reintroduzca como evaluación de un presente futuro. Esto no solo tiene consecuencias para la reproducción de la conocida brecha entre percepción de delito y sus cifras reales, sino que también tiene impactos negativos para la confianza futura en las instituciones de seguridad, en la evaluación de sus acciones y en las conductas de las personas. Es decir, a pesar de un mayor control del delito en el futuro, la percepción de descontrol actual mantendrá las conductas presentes aún si las condiciones cambian. Esta *histéresis* de las conductas individuales e institucionales requiere de intervenciones específicas especialmente en las comunas y zonas donde la percepción de inseguridad es mayor; intervenciones dirigidas tanto a los focos de producción de inseguridad delictiva como intervenciones urbanas y comunitarias para producir contextos espaciales que evoquen una mayor sensación de seguridad en la vida cotidiana.

Según las macrocategorías de delito que hemos construido (factores), el robo con violencia está principalmente concentrado en la Región Metropolitana, con algunos indicadores relevantes en comunas de la zona norte. Las agresiones y vandalismo se asocian con mayor intensidad a la zona centro sur y extremo sur; el robo a la vivienda en la zona centro sur; y los delitos económicos en la Región Me-

tropolitana, en la zona centro sur y, en menor medida, en el extremo sur. Independiente de su vinculación con eventos críticos, los distintos tipos de delito requieren intervenciones distintas. Este es un estándar conocido de la intervención sobre el crimen que se ha resuelto, por ejemplo, por militarización en la frontera norte para el control migratorio y también por militarización para el resguardo de infraestructura crítica en la zona sur. La militarización puede intervenir contextualmente en ese tipo de problemas, pero en un Estado de derecho la militarización no es una solución generalizable para el control del orden público. Para la desarticulación del delito son fundamentales los agentes y unidades especializadas de intervención directa en los sistemas criminales, la producción de conocimiento acerca de las dinámicas y fuentes del crimen, y la coordinación con las instituciones públicas y privadas que pueden contribuir a prevenir y mitigar los efectos del crimen.

El nuevo Ministerio de Seguridad Pública debe tener en cuenta que la seguridad no es solo capturar bandas, encerrar individuos o exhibir decomisos de drogas y armas. Para ser efectivo en el mediano plazo, debe manejar dos equilibrios delicados. El primero consiste en combinar la efectividad inmediata de sus estrategias de intervención delictiva y la coordinación de la intervención en los procesos sociales que producen el crimen y que no están bajo su competencia directa, sino a cargo de otras instituciones tanto públicas como privadas. El segundo consiste en el acto de balance entre la focalización de estrategias de intervención en zonas o comunas como las indicadas en este artículo y la mantención del orden público a nivel general del país. La seguridad pública no es un objeto de acción subsidiaria del Estado, sino su función más básica. Su focalización en zonas o comunas particulares, por tanto, siempre supone el riesgo de desproteger otras y, con ello, de no resguardar a otros grupos de ciudadanos, lo que es inaceptable desde un punto de vista normativo y político. Las dinámicas del crimen, especialmente del crimen organizado, captan estos vacíos con prontitud y se mueven a otras zonas donde continuar con sus actividades delictivas.

Ambos equilibrios deben funcionar en conjunto: los desplazamientos criminales territoriales producto de las focalizaciones de la intervención requieren de la coordinación de acciones directas sobre el delito y de la protección generalizada de instituciones sociales que previenen el ingreso al crimen en el largo plazo. Nada de esto es simple, es decir, no se resuelve con la creación de un ministerio, de más cárceles o con militarización. Estas son acciones necesarias, pero triviales. El problema está en observar las operaciones de fondo que constituyen la dinámica social, territorial y temporal del crimen y ajustar las operaciones de las instituciones y agentes encargados de su resolución a esos requerimientos.

Referencias

- Biderman, A. y Reiss, A.** 1967. On Exploring the 'Dark Figure' of Crime. *The ANNALS of the American Academy of Political and Social Science* 374(1), 1-15.
- Brantingham, P., Brantingham P. y Andresen, M.** 2017. The Geometry of Crime and Crime Pattern Theory (98-115). En Wortley, R. y Townsley, M. (eds.), *Environmental Criminology and Crime Analysis*. New York: Routledge.
- Bruinsma, G. y Johnson, S.** 2018. Environmental Criminology: Scope, History, and State of the Art (1-32). En Bruinsma, G. y Johnson, S. (eds.), *The Oxford Handbook of Environmental Criminology*. New York: Oxford University Press.
- Brunton-Smith, I., Cernat, A., Pina-Sánchez, J. y Buil-Gil, D.** 2024. Estimating the reliability of crime data in geographic areas. *The British Journal of Criminology* 64(6), 1347-1361.
- Buil-Gil, D., Medina, J. y Shlomo, N.** 2021. Measuring the Dark Figure of Crime in Geographic Areas: Small Area Estimation from the Crime Survey for England and Wales. *The British Journal of Criminology* 61(2), 364-388.
- Cadena-Urzúa, P., Iftimi, A. y Montes-Suay, F.** 2023. Seguridad ciudadana como elemento fundamental para el desarrollo humano: análisis de delitos en Chile mediante un modelo logístico espaciotemporal. *Revista Iberoamericana de Estudios de Desarrollo* 12(1), 30-49.
- ENUSC** 2023. Encuesta Nacional Urbana de Seguridad Ciudadana, ENUSC. Disponible en: <https://www.ine.gob.cl/estadisticas/sociales/seguridad-publica-y-justicia/seguridad-ciudadana> [11 de diciembre 2024].
- Felson, M.** 2017. The Routine Activity Approach (87-97). En Wortley, R. y Townsley, M. (eds.), *Environmental Criminology and Crime Analysis*. New York: Routledge.
- Gamarra, C., Rozas, J., Cardeiro, A. y Mascareño, A.** 2024. El lado oculto del delito. Aproximaciones al subreporte de la delincuencia en Chile. *Puntos de Referencia* 711, Centro de Estudios Públicos. Disponible en: <https://www.cepchile.cl/investigacion/el-lado-oculto-del-delito-aproximaciones-al-subreporte-de-la-delincuencia-en-chile/> [11 de diciembre 2024].
- Groff, E.** 2018. Do We Really Need Collective Social Process to Understand Why Crime Occurs and Offenders Commit Crime? (105-118). En Bruinsma, G. y Johnson, S. (eds.), *The Oxford Handbook of Environmental Criminology*. New York: Oxford University Press.
- Hastie, T., Tibshirani, R. y Friedman, J.** 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York: Springer.
- Hayward, K.** 2007. Situational Crime Prevention and its Discontents: Rational Choice Theory versus the 'Culture of Now'. *Social Policy & Administration* 41(3), 232-250.

- Kaufman, L. y Rousseeuw, P.** 2009. *Finding Groups in Data: An Introduction to Cluster Analysis*. New Jersey: John Wiley & Sons.
- Lohr, S.** 2019. *Measuring Crime: Behind the Statistics*. Boca Raton: CRC.
- Luneke, A.** 2016. Inseguridad ciudadana y diferenciación social en el nivel microbarrial: el caso del sector Santo Tomás, Santiago de Chile. *EURE* (42)125, 109-129.
- Mackles, A.** 2012. How to Get the Optimal K-Means Cluster Solution. *The Stata Journal* 12(2), 347-351.
- Navarrete-Hernández, P., Luneke, A. Truffello, R. y Fuentes L.** 2023. Planning for Fear of Crime Reduction: Assessing the Impact of Public Space Regeneration on Safety Perceptions in Deprived Neighborhoods. *Landscape and Urban Planning* 237(104809), 1-13.
- Riascos Villegas, Á., Ñungo, J., Gómez Tobón, L., Dulce Rubio, M. y Gómez, F.** 2023. Modelling Under-reported Spatio-temporal Crime Events. *PLOS ONE* 18(7), e0287776.
- Schwarz, A.** 2008. *Lokale Scoring-Modelle*. Lohmar: Eul Verlag.
- Skogan, W.G.** 1977. Dimensions of the Dark Figure of Unreported Crime. *Crime & Delinquency* 23(1), 41-50.
- Weinborn, C., Ariel, B., Sherman, L. y O'Dwyer, E.** 2017. Hotspots vs. Harmspots: Shifting the Focus from Counts to Harm in the Criminology of Place. *Applied Geography* 86, 226-244.
- Visauta, B. y Martori, J.** 2003. *Análisis estadístico con SPSS para Windows. Volumen II. Estadística multivariante*. Madrid: McGraw-Hill Interamericana.
- Xie, M. y Baumer, E.** 2019. Crime Victims' Decisions to Call the Police: Past Research and New Research. *Annual Review of Criminology* 2(1), 217-240.

Anexos

Estadística descriptiva de las variables a utilizar en el análisis factorial

La Tabla A1 muestra el promedio, desviación estándar, mínimo y máximo de cada variable. Las variables denotan la proporción de hogares que son victimizados a nivel comunal por cada categoría de delito. Esto es estimado a través de un método *Bootstrap* con muestras de 5.000 casos (con reemplazo), y 40 repeticiones —del mismo modo que Gamarra et al. (2024). Este proceso, y el hecho de que sean promedios comunales y no por hogar, puede resultar en que los valores promedio presentados en la Tabla A1 no sean *exactamente* los promedios nacionales que reporta la ENUSC (2023), aunque las diferencias son más bien nimias.

TABLA A1. Estadística descriptiva de las categorías de delito utilizadas para el análisis factorial, a nivel comunal

Variable	Promedio (%)	Desviación estándar	Mínimo	Máximo
Robo con violencia	2,48	2,13	0	9,25
Agresiones	1,07	0,77	0	3,36
Robo por sorpresa	2,35	2,19	0	11,14
Amenazas	3,51	1,72	0	9,78
Extorsión	0,03	0,41	0	2,62
Robo de vehículos	1,56	1,58	0	8,40
Robo en la vivienda	2,91	1,48	0,16	6,76
Robo desde vehículos	11,69	6,96	0,94	37,52
Hurto	4,67	2,41	0,37	13,13
Fraude bancario	4,57	2,22	0,54	14,59
Estafa	2,92	1,37	0	7,22
Soborno	0,07	0,18	0	1,29
Vandalismo de vehículos	7,45	3,84	0	22,91
Vandalismo en la vivienda	1,86	1,17	0	7,09
Bullying	0,08	0,60	0	2,43
Suplantación de identidad	0,07	0,66	0	2,97

Nota: N=134. Estimaciones a través de método Bootstrap con 40 repeticiones y muestras de 5.000 observaciones con reemplazo. Se presenta la proporción promedio de hogares victimizados, su desviación estándar, mínimo y máximo. La comuna es la unidad de análisis.

Fuente: Elaboración propia en base a datos ENUSC 2023.

Sobre el análisis factorial

El análisis factorial es una técnica estadística que permite la extracción de la mínima cantidad de factores que puedan explicar una matriz de muchas variables (16 en nuestro caso). En el caso del componente principal, esto se realiza a través de la combinación lineal de todas las variables (siendo el primer factor el que captura mayor cantidad de varianza total). Como anuncian Bisauta y Martori (2003), se busca reducir un conjunto de variables intercorrelacionadas a un número más pequeño de factores no relacionados. Adicionalmente, se aplicó un método de rotación ortogonal para que los factores no estén correlacionados entre ellos. En nuestro contexto, es ideal realizar la rotación debido a que utilizaremos los factores predichos para la generación de clústeres, por lo que cada dimensión debiese poseer información única.

En la Tabla A2 se presenta la proporción de varianza, de la varianza total, que es capturada por los 4 primeros factores seleccionados. Entre los 4 factores se captura aproximadamente 56% de la variabilidad total, lo que es un porcentaje aceptable.

TABLA A2. Varianza explicada de los 4 primeros factores, utilizando la victimización por 16 delitos consultados en la ENUSC 2023, a nivel comunal

Factor	Proporción de varianza explicada	Acumulada
Factor 1	0,2299	0,2299
Factor 2	0,1384	0,3683
Factor 3	0,104	0,4723
Factor 4	0,0826	0,5548

Nota: La varianza explicada es representada como la proporción de la variabilidad que es capturada por cada factor. Del total de factores calculados, se retienen 5 debido a que los demás poseen eigenvalues inferiores a 1 (criterio Kaiser). De los 5 factores restantes, se eligió restringir a los 4 primeros debido a que, de esta forma, cada factor de forma independiente es capaz de capturar al menos 8% de la variabilidad de los datos.

Fuente: Elaboración propia en base a datos ENUSC 2023.

En la Tabla A3 se muestran las relaciones entre los factores y las variables utilizadas en el análisis factorial. De esta tabla se puede distinguir qué variables están contenidas en cada dimensión o factor. Asimismo, también se presenta la proporción de la varianza que no es capaz de ser explicada por alguna comunalidad para cada variable (qué porcentaje de la varianza de cada delito no puede ser explicado por los 4 primeros factores, mientras mayor la varianza menor la importancia de dicha variable para la generación de los factores). Finalmente, para la predicción de los factores a nivel de la unidad de análisis se utilizó el método de *regression scoring*, que está basado en los parámetros presentados en la Tabla A4.

TABLA A3. Factor loadings del análisis de correspondencia ajustados por rotación ortogonal varimax

	Factor 1	Factor 2	Factor 3	Factor 4	Unicidad
Robo con violencia	0,8866	0,1409	-0,0376	-0,018	0,1923
Agresiones	0,3442	0,4595	0,2687	-0,0398	0,5641
Robo por sorpresa	0,8894	0,0557	0,0572	0,0974	0,1924
Amenazas	0,086	0,741	0,0739	0,3001	0,3429
Extorsión	0,1364	0,0422	0,0247	0,8466	0,2543
Robo de vehículos	0,7601	0,1208	0,0439	-0,0052	0,3792
Robo en la vivienda	0,0277	0,2394	-0,3476	0,3938	0,4614
Robo desde vehículos	0,8052	0,1399	0,2732	0,0508	0,2511
Hurto	0,6056	0,1888	0,2065	0,3553	0,4283
Fraude bancario	0,4078	-0,0976	0,7237	0,1274	0,2841
Estafa	0,01	0,2293	0,7384	-0,0664	0,3679
Soborno	-0,2284	0,1953	0,3073	0,3774	0,632
Vandalismo de vehículos	0,3177	0,6847	0,129	-0,1265	0,3775
Vandalismo en la vivienda	0,1513	0,7595	-0,0502	0,0151	0,3617
Bullying	0,0507	0,3894	0,3778	0,1796	0,5649
Suplantación de identidad	0,1349	0,0687	0,1253	-0,0859	0,2395

Nota: Las columnas muestran las correlaciones entre los factores subyacentes del análisis factorial y las variables utilizadas (16 delitos). Una elevada correlación, ya sea positiva o negativa, indicaría que dicho delito se ve representado en dicho factor (se destacan los valores más relevantes). La última columna, Unicidad, indica qué tan única es la información que posee dicha variable. Es decir, cuánto de la variación en esa variable no es explicada por los factores 1, 2, 3 y 4. Mientras mayor unicidad, mayor será la información única que posea dicha variable.

Fuente: Elaboración propia en base a datos ENUSC 2023.

TABLA A4. Scoring coefficients, método de regresión basado en los factores con rotación varimax

Variable	Factor 1	Factor 2	Factor 3	Factor 4
Robo con violencia	0,2874	-0,00931	-0,14815	-0,07254
Agresiones	0,04421	0,21486	0,08374	-0,12462
Robo por sorpresa	0,27988	-0,08822	-0,07365	0,03248
Amenazas	-0,07013	0,36769	-0,05753	0,13356
Extorsión	0,00574	-0,10565	-0,03063	0,68006
Robo de vehículos	0,22869	-0,03359	-0,06503	-0,05146
Robo en la vivienda	-0,01524	0,07613	-0,25957	0,30897
Robo desde vehículos	0,21699	-0,05317	0,08063	-0,01867
Hurto	0,14139	-0,03529	0,04197	0,2337
Fraude bancario	0,05603	-0,20244	0,46501	0,08187
Estafa	-0,11545	0,03577	0,49925	-0,10192
Soborno	-0,16116	0,02738	0,22057	0,2895
Vandalismo de vehículos	0,0151	0,3447	-0,01728	-0,21456
Vandalismo en la vivienda	-0,03606	0,40194	-0,13252	-0,0958
Bullying	-0,06192	0,17641	0,1896	0,0716
Suplantación de identidad	-0,01591	-0,05687	0,11241	-0,06703

Sobre el número de clústeres óptimos

Para encontrar el número óptimo de clústeres, se utiliza la propuesta de Makles (2012). En esta se analiza los siguientes estadísticos: la suma *wilthin* de los cuadrados (WSS, también su logaritmo $\log(WSS)$), el coeficiente η^2 y la reducción proporcional del error (PRE) (Makles 2012), para un posible conjunto de K soluciones. En términos matemáticos, WSS es definido de acuerdo con la ecuación [A1]. Donde k es el clúster dentro del conjunto de K clústeres, S_k representa el conjunto de puntos en el clúster k y X indica es un punto dentro del clúster S_k . El parámetro μ_k es el centroide del clúster k . Entonces, la fórmula dentro de las sumatorias calcula la distancia al cuadrado entre el punto X y el centroide, lo que brinda una idea de cuán lejos está cada punto de su centro. Esto se repite para todos los puntos pertenecientes el clúster, y para todos los clústeres k dentro del conjunto K de posibles soluciones.

$$WSS = \sum_{k=1}^K \sum_{x \in S_k} |X - \mu_k|^2 \quad [A1]$$

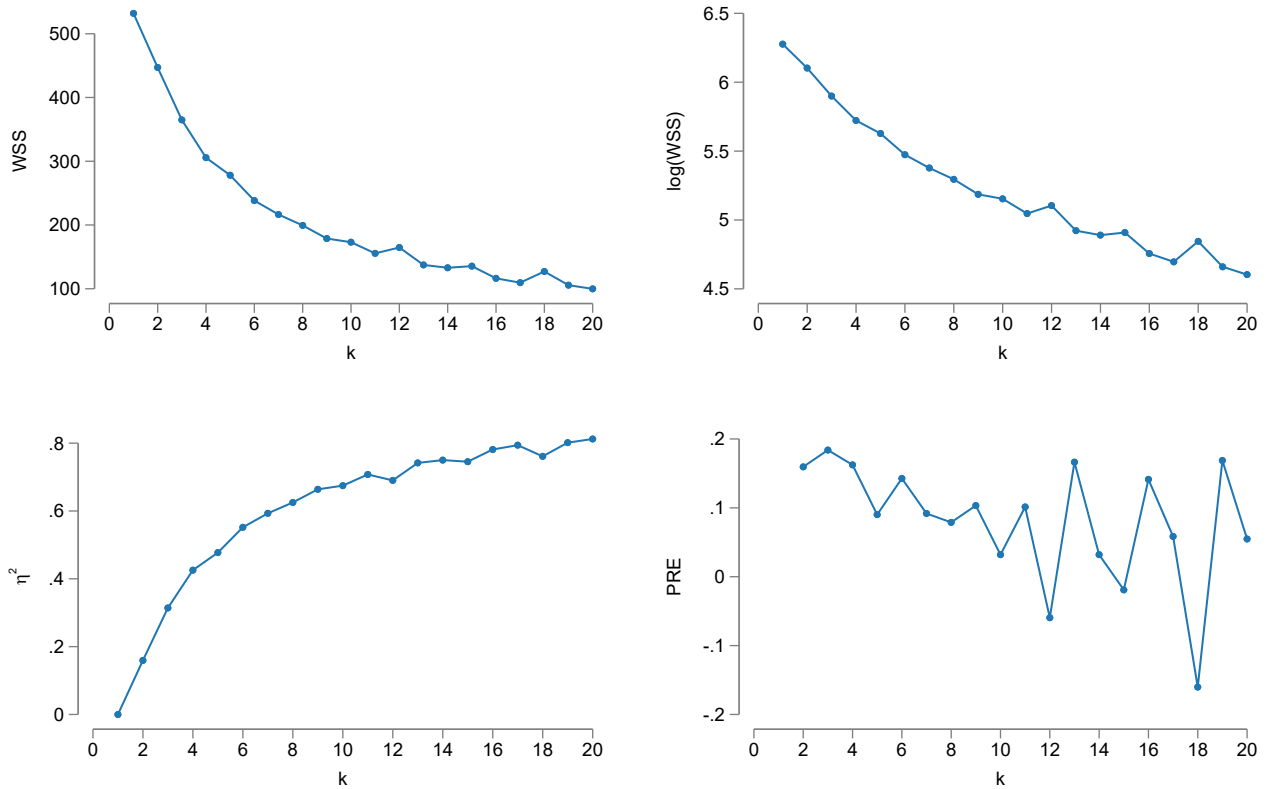
Para complementar la métrica anterior, también se calcula el coeficiente η^2 y la reducción proporcional del error (PRE). Ambos estadígrafos pueden ser definidos de acuerdo con [A2] y [A3] (Schwarz 2008). El coeficiente η^2 indica la reducción proporcional de los WSS para cada clúster k comparada con la suma total de cuadrados (TSS). Por otro lado, PRE representa la reducción proporcional de los WSS para el clúster k comparado con la de $k-1$. Es decir, mide cuánto más reduce la solución k con respecto a la solución previa (Makles 2012; Schwarz 2008). $WSS(1)$ es igual a la suma total de cuadrados debido a que es el WSS cuando el clúster de los datos es igual a 1, es decir, la data no se agrupa en subdivisiones y se considera en su totalidad.

$$\eta_k^2 = 1 - \frac{WSS(k)}{WSS(1)} = 1 - \frac{WSS(k)}{TSS} \quad \forall k \in K \quad [A2]$$

$$PRE_k = \frac{WSS(k-1) - WSS(k)}{WSS(k-1)} \quad \forall k \in K \quad [A3]$$

Estos estadísticos se presentan para todas las posibles soluciones k testeadas ($K=20$) en la Figura A1. De esta figura, lo primero que es ostensible es que la solución óptima no es tan evidente como en otros casos, como el ejemplo presentado por Makles (2012). Es decir, la reducción en la suma de cuadrados no se aproxima fuertemente a 0 a partir de algún valor (solución) k , ni tampoco el coeficiente η^2 posee una pendiente plana a partir de un valor (solución) k . Sin embargo, se puede inferir que, para $k=6$, existe un cambio en la reducción de cuadrados presente en los paneles superiores de la Figura A1 (es decir, la pendiente se vuelve ligeramente más plana a partir de ese valor). Además, en cuanto al parámetro PRE, se observa que la reducción en los cuadrados de $k=6$ con respecto a $k=5$ es elevada y que a partir del $k=6$ en adelante dicho valor decrece o es altamente volátil. Entonces, se interpreta que el $k=6$ es la solución óptima.

FIGURA A1. Estadísticos para discernir el número óptimo de clústeres

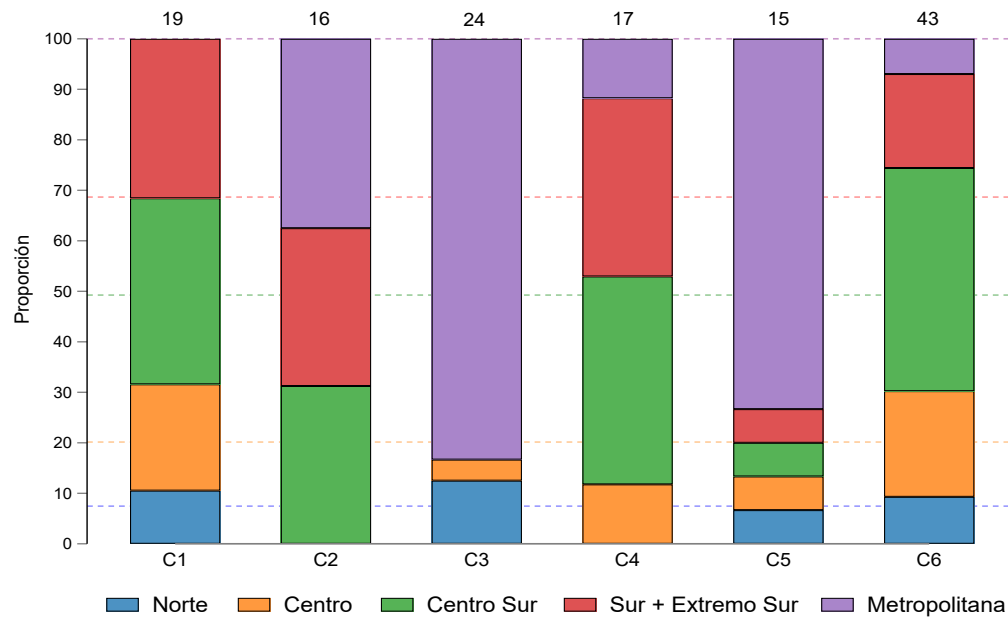


Nota: El panel superior izquierdo presenta la suma interna de cuadrados, ecuación [1]. El panel superior derecho es el logaritmo de la ecuación [1]. El panel inferior izquierdo presenta el coeficiente η^2 y el panel inferior derecho la reducción proporcional del error (PRE). Se muestran los estadígrafos para las primeras 20 posibles agrupaciones utilizando el algoritmo k-medias.

Representación de macrozonas en cada clúster

La Figura A2 muestra la distribución de las comunas en cada clúster, distinguiendo a qué macrozona representan. Se observa que el C1 sobrerrepresenta al Norte, Centro, Centro sur y al Sur + Extremo sur. El C2 y C4 sobrerrepresentan a las macrozonas Centro sur y Sur + Extremo sur. El C3, por su parte, sobrerrepresenta únicamente a la región Metropolitana y a la macrozona Norte, mientras el C5 únicamente a la región Metropolitana. Finalmente, el C6 sobrerrepresenta a las macrozonas Centro y Centro sur.

FIGURA A2. Representación (%) de las macrozonas dentro de cada clúster



Nota: El número de comunas en cada clúster se muestra encima de las barras. Para una mejor lectura del gráfico, las proporciones que abarca cada macrozona son las siguientes: 7,5% (Norte), 12,7% (Centro), 29,1% (Centro sur), 19,4% (Sur + Extremo sur) y 31,3% (Metropolitana). Una forma intuitiva de interpretar los valores mostrados en la figura es observar si la proporción dentro de un clúster que abarca una determinada macrozona es inferior (o superior) a la proporción que dicha macrozona representa en el conjunto total de datos, esa macrozona está subrepresentada (sobrerrepresentada) en ese clúster. Por ejemplo, la Metropolitana está sobrerrepresentada en C3, pero subrepresentada en C6.



CENTRO DE ESTUDIOS PÚBLICOS



Cada artículo es responsabilidad de su autor y no refleja necesariamente la opinión del CEP.

Director: Leonidas Montes L.
Coordinador académico: Sebastián Izquierdo
Diagramación: Pedro Sepúlveda V.



[VER EDICIONES ANTERIORES](#)

