



INSTITUTO DE ESTUDIOS URBANOS Y TERRITORIALES
FACULTAD DE ARQUITECTURA, DISEÑO Y ESTUDIOS URBANOS

LOS ESPACIOS Y ESCALAS DE LA SEGREGACIÓN RESIDENCIAL EN SANTIAGO DE CHILE (2012)

Un enfoque desde sus geografías locales y nuevas directrices para
resolver su escala de análisis

Tesis presentada para obtener el grado académico de Magíster en Desarrollo Urbano

Gabriel Acuña

Profesor guía: Ricardo Truffello

Instituto de Estudios Urbanos y Territoriales

Pontificia Universidad Católica de Chile

30 de Julio de 2019

RESUMEN

En el marco de los estudios sobre segregación residencial, el debate sobre cómo medirla sigue organizando buena parte de la discusión. Dicha problemática metodológica se aloja en torno a la insensibilidad que los índices sintéticos más populares presentan frente a la distribución y agregación espacial de los datos; respectivamente, se conocen estos como “el problema del tablero de ajedrez” y “el problema de la unidad espacial modificable”. Aun cuando existen diversas soluciones operativas, estas no se hacen cargo de cierto trasfondo teórico particular: la escala de análisis de la segregación residencial.

Considerando la escala geográfica como categoría epistemológica, analiza esta investigación un método para determinar la escala óptima para medir segregación residencial. A través de la operacionalización local de tres índices de segregación (Disimilaridad, Aislamiento y Theil), se comparan resultados de aplicar rutinas de clustrización jerárquica a cada uno de esos índices, relevando la relación subyacente entre el índice medido (dimensión particular de segregación) y su escala óptima de análisis.

ÍNDICE

CAPÍTULO I. INTRODUCCIÓN	4
1. La escala como cuestión epistemológica: rompiendo las trabas del debate ...	5
2. Pregunta de investigación	11
3. Objetivos.....	11
4. Hipótesis.....	12
5. Aproximación metodológica	12
CAPÍTULO II. SEPARANDO LA PAJA DEL TRIGO	17
1. Un siglo de estudios sobre SR: enfoques, aportes y lecciones.....	18
2. Medición de segregación en Latinoamérica:.....	22
3. Debate metodológico: separando la paja del trigo	23
CAPÍTULO III. METODOLOGÍA DE LA INVESTIGACIÓN	26
1. Índices espaciales de SR: capturando sus principales dimensiones en el entorno local.....	27
2. Evaluando la escala mediante un algoritmo de regionalización jerárquica	33
CAPÍTULO IV. ESPACIO Y ESCALAS DE LA SEGREGACIÓN RESIDENCIAL EN SANTIAGO	36
1. El patrón de segregación residencial en el Gran Santiago.....	37
2. Evaluando las escalas de la SR.....	46
3. Conclusiones y nuevas directrices de investigación.....	53
4. Reflexión final: segregación residencial y su escala	55
REFERENCIAS BIBLIOGRÁFICAS	57

CAPÍTULO I.

INTRODUCCIÓN

1. La escala como cuestión epistemológica: rompiendo las trabas del debate

En el marco de los estudios sobre segregación residencial –de ahora en adelante, SR–, el debate sobre cómo medirla sigue organizando buena parte de la discusión. Luego, es menester aclarar que se entiende por SR.

En una de las más influyentes y citadas publicaciones que existe sobre SR, Massey y Denton (1988) la definen como “el grado en el cual dos o más grupos viven separados uno del otro, en distintos lugares de entorno urbano” (p.282). Algunos años antes, White (1983) distingue entre aquella segregación sociológica, que encuentra su significado en la “ausencia de interacción entre grupos sociales”, de su símil geográfico, que entiende el fenómeno en cuanto a la “dispar distribución de los grupos sociales a lo largo de espacio físico” (p.1009). Esta última aparece más en línea con la definición de Massey y Denton (1988).

Marcuse (2005), en torno a la necesidad de usar el término “segregación” con la intención de referirse a “patrones que son indeseables”, la describe como “el proceso por el cual un grupo de población tratada como inferior es forzada, vale decir involuntariamente, a agruparse en un área espacial definida, en un gueto” (p.16).

Pivotando sobre la visión de Marcuse en cuanto corresponde a procesos de “cierre, exclusión y control social” (Tilly, 2000: 21) y recalcando su carácter incuestionablemente espacial, se comprenderá como SR a la concentración espacial homogénea que, por su forma segregada de habitar el espacio urbano, refuerza efectos de desintegración social en detrimento de la calidad de vida de las clases sociales vulnerables.

En relación a esa espacialidad que se debe tener presente a la hora de estudiar la SR relevo, antes de abordar las distintas posturas respecto a su medición, el papel crucial de la escala geográfica en la cual se manifiesta y la necesidad de considerarla como categoría epistemológica (Jones, 1998). La razón principal de considerar la escala como una “*manera de conocer o aprehender*” (p.28), tiene directa relación con pensar cada configuración socio-espacial de SR en su propia

espacialidad, entendiendo de esta última que *“no son los fenómenos sociales [los que ocurren] en el espacio sino ambos fenómenos sociales y espacio los constituidos a partir de las relaciones sociales”* (Massey, 1994: 2). Mientras la espacialidad de la SR se concibe inherente en su contexto, es posible comparar su escala –como categoría epistemológica– con el nivel de observación de cualquier fenómeno a analizar por parte de cualquier ciencia en particular. Pertinente sería preguntarse, como lo hace Wagensberg (2004) sobre *“el interés de considerar más de un nivel de observación de una misma realidad [siendo que] cada nivel de observación tiene su propia inteligibilidad”* (p. 48).

Paso a explicar la decisión de enfrentar este debate utilizando los conceptos de espacialidad y escala.

Sobre la discusión de cómo medir la SR

Dos publicaciones realizadas por investigadores latinoamericanos relevan, a través de una buena revisión bibliográfica, la idoneidad de migrar de los índices de medición de segregación más utilizados –disimilaridad, aislamiento e interacción– hacia la aplicación de unos basados en la auto-correlación espacial (Garrocho y Campos-Alanís, 2013; Ruiz-Tagle y López, 2014). En cuanto a las investigaciones provenientes del norte global, aparecen novedosas modificaciones a índices convencionales no-espaciales (Lee y Culhane, 1998; Wong, 2003; Reardon y Firebaugh, 2002; Reardon y O’Sullivan, 2004; Monkkonen y Zhang, 2014). Independiente de los enfoques y contextos urbanos particulares de estos artículos, coinciden los autores en cuestionar las pesquisas basadas en índices convencionales para medir la SR. Más allá de las falencias operacionales que deben afrontar las distintas propuestas metodológicas para medir la SR (Openshaw, 1983; White, 1983; Bailey y Gatrell, 1995; Wong, 2004), subyace tras esta controversia dos aspectos cruciales.

El primero, una profunda discusión sobre el enfoque teórico para explicar el fenómeno de SR socioeconómica y la interpretación de sus resultados. Emergen,

por una parte, cuestiones sobre la distinción entre las desigualdades sociales y geográficas (White, 1983) y su crítica hacia el correlato entre ellas, definida por Sabatini, Cáceres y Cerda (2001) como la tesis del espejo. Estrecha relación existe entre estos postulados y aquellas investigaciones basadas en dos paradigmas sobre las consecuencias de la SR y su conexión con problemas sociales específicos de desintegración social: el efecto barrio (Jencks y Mayer, 1990; Sampson, 2012) y la geografía de oportunidades (Galster y Killen, 1995), como las más aludidas. En la perspectiva de sus críticos, se le reprocha a este enfoque desestimar la acción de otros procesos sociales más allá de la expresión física de distancia (o proximidad) de las clases sociales (Ruiz-Tagle, 2016), la retirada del Estado del Bienestar (Bourdieu, 1999; Wacquant, 2009), la relación entre proximidad territorial de distintas clases sociales en barrios cerrados y la estigmatización (Caldeira, 2007; Wacquant, 2007) y de un empirismo que corre el peligro de caer en problemas de reificación del espacio (Castells, 1977; Bourdieu, 1999).

El segundo, muchas veces soslayado por considerarse simple dato del estado del arte, sobre la teorización en torno a las dimensiones socio-espaciales que deben estar presentes a la hora de estudiar y medir SR. Ante la necesidad de discutir y operacionalizar la SR, Massey y Denton (1988) conciben que, al tratarse de un fenómeno que deriva de “una compleja interacción de diferentes procesos sociales y económicos” en la estructura espacial urbana (p.309), se compone de cinco dimensiones socio-espaciales. Sabatini *et al* (2001) consideran que sólo dos de estas son relevantes, la de *uniformidad* y la de *exposición*, por medir distintas formas objetivas de habitar el espacio por los grupos sociales. De similar forma, Reardon y O’Sullivan (2004) llegan a la misma conclusión: mientras la uniformidad mide el grado de “similitud en cuanto a la distribución de los grupos en el espacio residencial”, la exposición se relaciona a la “experiencia del entorno por los individuos” (p.125). Para efectos de esta investigación, alrededor de esas dos dimensiones se desarrollará el estudio de los índices de SR y su escala geográfica.

Distintas aproximaciones al concepto de escala geográfica

Desde la arena de la geografía urbana, la discusión teórica sobre el concepto de la escala se intensificó a principios de la década del 2000 (Brenner, 2000, 2001; Marston, 2000; Marston y Smith, 2001). Para Brenner (2000) las escalas geográficas son una construcción social, susceptibles de reordenación y diferenciación jerárquica según “relaciones, procesos y dinámicas socio-espaciales” específicas en el contexto de las estructuras del capitalismo (Brenner, 2001: 592). Las críticas de Marston (2000) y Marston y Smith (2001) a los postulados de Brenner apuntan a su falta de consideración de las interacciones sociales a nivel local (Marston y Smith, 2001). Cabe agregar que, en su crítica inicial, Marston le fustiga a Brenner la ausencia de una explicación que considere “otros sistemas de dominación”, más allá de la tradicional causalidad estructural asociada al sistema capitalista –dando como ejemplo los estudios desde el feminismo a las relaciones sociales patriarcales– (2000: 219).

Si bien la discusión sobre la escala en los estudios sobre SR ha tenido un tratamiento preferentemente operacional, algunos autores han llegado a conclusiones similares a aquellas ligadas a la construcción social del espacio y sus escalas (Massey, 1994, 2005; Marston, Jones y Woodward, 2005). Para esta corriente, la escala es vista como área de análisis o extensión territorial, en relación al ámbito metodológico y los problemas que suscita la unidad espacial modificable –desde ahora MAUP, por sus siglas en inglés– (Openshaw, 1983; Wong, 2004). En su análisis, Openshaw (1983) revisa que el problema de la escala tiene dos dimensiones: la extensión y la forma (geometría) de agregación de datos. Al proponer un procedimiento iterativo para agregar unidades espaciales en nuevas zonas de análisis para reducir el MAUP, Openshaw intuye la necesidad teórica de determinar previamente los fenómenos que se quieren predecir, y que estos, por su especificidad, implicarían distintas escalas y formas de agrupar las unidades espaciales en cuestión. Asimismo, Sabatini *et al* (2001) y Rodríguez (2001) resaltan la importancia de determinar previamente la escala de análisis, dependiendo del fenómeno socio-territorial que se quiera medir.

Justificación

Hasta ahora he apuntado diversas consideraciones en torno a la conveniencia de cambiar la forma de comprender aquellos fenómenos socio-espaciales relacionados con los estudios de SR. Visto el concepto de escala alrededor de sus declinaciones teóricas y metodológicas, debe repensarse la espacialidad como dimensión ontológica –con su correlato epistemológico en la escala– de los distintos fenómenos y procesos en el espacio (Massey, 1994, 2005; Soja, 2010), desnaturalizando las posibles consecuencias de la SR. Desde una perspectiva similar, sobre la expresión física de las injusticias en el espacio urbano y sus escalas geográficas, Soja (2010) sostiene:

(...) estas escalas de resolución geográfica no debiesen ser vistas como capas discretas inconexas y naturales de origen. Se interconectan entre ellas y, como la espacialidad misma, están socialmente producidas y abiertas a ser reestructuradas y reorganizadas. (p. 213).

Aprehendida dicha perspectiva, atiende mi interés por reflexionar sobre la epistemología en los estudios de SR dos cuestiones fundamentales: primero, porque es *“tarea de la práctica espacial teóricamente informada el relevar aquellas estructuras de privilegio (...) a una superficie de conocimiento público y general”* (Soja, 2010: 56), en relación a la eficacia de las políticas públicas que pueden derivar de ello. En segundo lugar, al *“acercamiento entre las partes que no son capaces de discernir frente a un análisis que va más allá de su comprensión [a través de un] programa metodológico para crear una aproximación más sintética (...), integrando múltiples aspectos del proyecto crítico”* (Purcell, 2003: 317). Desde esa perspectiva, el estudio de los datos censales del Santiago de 2012 puede otorgar evidencia esclarecedora sobre los distintos procesos de segregación que ocurren en el entorno urbano en relación a su escala. Merece la pena entonces el volver la mirada a la controversia inicial sobre la idoneidad de los distintos índices sintéticos que existen para medir la SR.

Sostengo, pues, la necesidad de identificar y relevar la relación que existe entre cada índice y las escalas geográficas donde mejor se expresan (Sabatini, Wormald,

Sierralta y Peters, 2010), ya sea presente a sus formas estructurales que conforman los posibles espacios de la SR, como sus consecuencias medidas respecto a dinámicas relevantes en la producción y refuerzo de las injusticias locales en la ciudad.

2. Pregunta de investigación

Con base en la revisión bibliográfica del debate teórico, y bajo la premisa de que las escalas de análisis pertenecen a una categoría epistemológica esencial para los estudios de SR ¿Qué índices sintéticos miden de manera eficaz la SR en relación a sus dimensiones objetivas y la escala geográfica específica en que se presentan en el espacio?

3. Objetivos

3.1. Objetivo general

Analizar la relación que existe entre la segregación residencial, manifestada a distintas escalas geográficas, sus dimensiones socio-espaciales objetivas –eso es, uniformidad y exposición– y los distintos índices sintéticos para medirla.

3.2. Objetivos específicos

- Estudiar las distintas formas y escalas geográficas que toma la SR en el Gran Santiago del 2012 respecto a sus dimensiones de uniformidad –concentración espacial– y exposición –grado de homogeneidad–.
- Revisar y discutir los supuestos teóricos tras cada índice de segregación en los resultados de las mediciones realizadas.
- Profundizar el análisis en torno a la escala de análisis de la SR a través de un método de regionalización jerárquica y proponer directrices metodológicas para encontrar su nivel óptimo de análisis.

4. Hipótesis

Frente a las distintas construcciones socio-espaciales posibles, las alternativas de análisis respecto a la diferenciación social en el espacio varían según las definiciones relacionales sobre el fenómeno de SR y las dinámicas socio-territoriales que se presentarán desde su contexto histórico y local. Afirmando, similar a lo que señalan Sabatini *et al* (2010), que existen procesos de SR que se expresarán en el espacio con un alcance inter-barrial, y otros que son más propios de dinámicas que varían su alcance desde el ámbito comunal hasta el metropolitano. A su vez, los supuestos teóricos tras los distintos índices de SR relevan distintas dimensiones de este fenómeno. Por ello, índices que privilegien modelos de auto-correlación espacial lograrán mejores resultados para estimar la interacción y cobertura de las escalas geográficas que toman los procesos de SR, en cuanto relevan la contigüidad de la segregación como fenómeno socio-espacial. En cambio, los índices convencionales –disimilaridad, aislamiento– y sus derivados tendrán la ventaja a la hora de medir SR en escalas geográficas de mayor intensidad y extensión espacial, por cuanto el peso explicativo de la homogeneidad y concentración de la población segregada trasciendan a posibles discontinuidades del fenómeno.

5. Aproximación metodológica

5.1. Enfoque y técnicas de investigación

Escogido un enfoque epistemológico que releva la construcción social del espacio, esta investigación aborda el fenómeno de SR desde una perspectiva relacional entre sus distintas dimensiones socio-espaciales –cuantitativamente medibles a través de indicadores sintéticos de segregación– y su escala de análisis. Bajo dicha postura, y dado el carácter exploratorio tras las preguntas de investigación, la aproximación a este problema es principalmente deductiva (Hernández, Fernández y Baptista, 2010). Para dar cuenta del lugar que deben tomar las escalas geográficas dentro del abordaje escogido, el estudio se estructura en dos etapas.

En primer lugar, una acuciosa revisión de la literatura disponible sobre las dos corrientes teóricas que comparecen en la problematización de esta investigación: aquella sobre las teorías emanadas del fenómeno de SR y la evolución de las técnicas metodológicas para medirla (capítulo III); y la literatura correspondiente a la comprensión de la producción del espacio orientada desde la geografía humana (capítulo II). De este trabajo bibliográfico, pretendo examinar aquellos aspectos teóricos y metodológicos que pueden dar cuenta de nuevas perspectivas para el análisis de la SR –la cual se encuentra profundamente arraigada en su marco teórico al campo de la sociología urbana–.

La siguiente etapa de la investigación consiste de un análisis cuantitativo para determinar la configuración socio-espacial del caso estudiado respecto a tres factores críticos: el índice sintético de medición escogido (a); la escala geográfica de análisis, entendida como la extensión territorial (b); y el nivel de agregación de los datos (c). En ese sentido, los métodos cuantitativos de análisis contemplados son de alcance descriptivo para determinar la estratificación social y su configuración espacial, y de carácter relacional en su análisis sobre los índices, las escalas y el nivel de agregación de datos (Hernández *et al*, 2010). Con ese fin, dos técnicas de investigación cuantitativa son aplicadas.

En orden secuencial, la primera técnica consiste en calcular el valor que toman tres índices espaciales de segregación a nivel de entorno local para cada manzana del Gran Santiago –GS–: disimilaridad (D_{x_r}), aislamiento (xPx_r) y el índice de la teoría de información de Theil (H_r), adaptados del trabajo realizado por Reardon y O’Sullivan (2004). Sin embargo, no es de interés de esta tesis calcular el valor que toma cada índice a nivel metropolitano –o cualquier escala intermedia–, puesto que no está dentro del alcance de la investigación el probar su calidad como indicadores de SR; en cambio, sí lo es obtener patrones válidos de segregación en el GS según las dimensiones socio-espaciales que cada índice en particular mide.

Inspirada en la metodología desarrollada por Garretón y Sánchez (2016) para determinar escalas óptimas de análisis espacial, la segunda técnica toma como insumo los índices de segregación residencial previamente calculados y les aplica

un algoritmo de regionalización jerárquica-local. Los resultados de este método son lo que finalmente serán interpretados a la luz del análisis teórico de esta investigación, tanto en su distribución y escala geográfica como en su sensibilidad estadística para distintos niveles de clusterización.

5.2. Fuente y procesamiento de datos

El trabajo de investigación cuantitativo se basa en el procesamiento y análisis del índice socio material territorial –por sus siglas ISMT–, desarrollado por el Observatorio de Ciudades UC (OCUC, 2018) en base a los datos censales del año 2012 en el área metropolitana de Santiago. Para su cálculo, el ISMT pondera tres índices socio-materiales con especificidad territorial mediante análisis de escalamiento óptimo –también conocido como *Homals*–: escolaridad del jefe de hogar, nivel de hacinamiento en la vivienda y calidad material de esta misma.

Considerados los datos a nivel de manzana de 35 comunas contiguas dentro del área urbana consolidada de la ciudad de Santiago –las 32 comunas de la provincia de Santiago, Puente Alto, Lampa y San Bernardo–, se sirve la investigación del ISMT como variable de estratificación socioeconómica para el cálculo de los tres índices espaciales de segregación antes mencionados. A su vez, el puntaje continuo y normalizado del ISMT se utiliza para evaluar y escoger tres comunas del GS donde ejecutar el algoritmo de clusterización. Para validar que estas tres comunas sean tipológicamente distintas entre ellas respecto a su distribución social, se discrimina su nivel de autocorrelación espacial a través del índice global de Moran (Bailey y Gatrell, 1995).

Las herramientas computacionales utilizadas para los distintos procesos son:

- *ArcGis* (versión 10.5.1.; ESRI, 2017), para extraer pesos espaciales y relaciones de proximidad entre unidades espaciales, evaluar el índice global de Moran para el GS y sus comunas, medir estadísticas descriptivas y generar mapas.

- *RStudio* (versión 1.1.463; R Core Team, 2018) junto a sus paquetes *tidyverse* (versión 1.2.1; Wickham, 2017), *foreign* (versión 0.8-70; R Core Team, 2017) y *glue* (versión 1.3.0; Hester, 2018) para el cálculo de los índices espaciales de SR.
- *GeoDa* (versión 1.12.1.161; Anselin, Syabri y Kho, 2006), para la ejecución del algoritmo de regionalización jerárquica.

5.3. Justificación y limitaciones de la metodología propuesta

Pese a lo exhaustivo del estudio cuantitativo, asoman limitaciones generales que merecen ser explicitadas, para evitar cualquier mala interpretación. Las más específicas, en relación al procesamiento de índices de segregación y al método de clusterización, serán justificadas debidamente en el capítulo IV, posterior al debate sobre las técnicas de medición.

La primera advertencia, en relación al mismo trabajo cuantitativo, se desprende del alcance del análisis de datos. Al asumir una cantidad fija de dimensiones e índices, las conclusiones posibles tienden más a abrir posibilidades de estudios futuros que generar conocimiento conclusivo sobre la premisa de la escala de análisis como categoría epistemológica; en ese sentido, quedan fuera de análisis otras dimensiones de la SR, como puede ser las distancias absolutas –reales– a centralidades productivas o la movilidad dentro del territorio metropolitano.

A su vez, para contar con un nivel de desagregación de los datos a nivel de manzana para respetar de mejor manera el comportamiento a escala barrial de la SR, se tuvo que optar por utilizar datos que provienen del censo 2012. Aun cuando existe cálculo del ISMT respecto del censo 2017, estos están a nivel de zona censal, lo cual guarda limitaciones en relación a la morfología en que estos datos se agrupan.

Respecto a la decisión de usar el ISMT como *proxy* de estratificación socioeconómica, a pesar de los insoslayables inconvenientes económicos y sociológicos que padece al provenir de metodologías de estratificación de *marketing* (Ruiz-Tagle y López, 2014), esta la tomé de manera pragmática. La alternativa de

primera instancia fue intentar replicar la metodología que Agostini, Hojman, Román y Valenzuela (2016) utilizan para imputar ingresos del hogar a datos censales, pero a poco andar encontré falencias y dificultades que merecen ser estudiadas en una investigación aparte.

En la práctica, el factor de expansión calculado a nivel comunal por la Encuesta de Caracterización Socioeconómica Nacional –CASEN– no es representativo, por lo que no es válido utilizar las comunas como *cluster* de factor covariante para modelar el error; y aún cuando fuese válido, Elbers, Lanjouw y Lanjouw (2003: 362) –referencia de la investigación de Agostini *et al* (2016)– advierten que el error crece exponencialmente al simular los ingresos en unidades espaciales menores al tamaño de los *cluster* covariantes. Por cierto, esto no impide buscar un modelo alternativo de ingreso siguiendo los mismos lineamientos metodológicos; sin embargo, se encontraron indicios de que esto sólo sería posible mediante la aplicación de modelos de regresión no lineales más complejos.

Finalmente, es relevante considerar el peso relativo que tienen las dimensiones subjetivas sobre el habitar segregado de las distintas clases sociales en el entorno urbano –percepción de la segregación–. A la hora de esgrimir un enfoque basado en la construcción social del espacio, queda pendiente la verificación de los supuestos teóricos y resultados del análisis cuantitativo el correlato cualitativo sobre el espacio urbano estudiado.

CAPÍTULO II.

SEPARANDO LA PAJA DEL TRIGO

**CONSIDERACIONES TEÓRICAS Y METODOLÓGICAS PARA
INVESTIGAR LA SEGREGACIÓN RESIDENCIAL**

No es de extrañar que los enfoques predominantes en los estudios sobre SR abarquen desde procesos globales a locales de transformación social, económica y espacial, dada la “interacción de influencias endógenas y exógenas” sobre la producción de “geografías de discriminación” en el espacio urbano (Soja: 2010: 55). Surge de esto la pregunta sobre qué elementos y dimensiones incorporar a la hora de medir SR, donde la coherencia metodológica parece alojarse en los motivos investigativos de medir segregación –¿Por qué medir la SR?–. Con el fin de distinguir los distintos elementos que deben ser tomados en cuenta para el debate metodológico en torno a la medición de segregación, se analizan a continuación: los aportes históricos desde distintas corrientes teóricas (i); y aquellos aspectos conceptuales específicos de la experiencia latinoamericana que son necesarios precisar (ii).

1. Un siglo de estudios sobre SR: enfoques, aportes y lecciones

Al retroceder un siglo para revisar estudios de SR, no quiero decir bajo ninguna circunstancia que previamente no se hablara sobre dicho fenómeno en el espacio urbano. Ejemplos de esto son los estudios y visiones sobre la ciudad industrial del siglo XIX que recoge Soja (2008) de los estudios de Engels sobre Manchester en 1844, como aquel paseo en forma de panfleto que Hall (1996) evoca sobre la experiencia de Mearns en el Londres marginado de 1883. La razón para fijar el cero temporal a principios del siglo XX considera dos factores: el primero, por lo que significa la escuela sociológica de Chicago en cuanto al estudio sistemático sobre SR. El segundo, en relación a las críticas que recibe el enfoque basado en la ecología urbana. Comenzar desde este punto permite analizar mejor las distintas posturas teóricas que hoy día prevalecen –y divergen– al tratar la relación entre SR y fenómenos de desintegración socio-espacial de interés para la toma de decisiones públicas.

Como sugiere el nombre adoptado por la escuela de Chicago, el enfoque ecológico para aproximarse al estudio socio-espacial de la ciudad asume en esta un comportamiento orgánico, producto de la naturaleza humana de quienes integran la

ciudad. Para Park, en la ciudad convergen fuerzas que del agrupamiento humano e institucional producen un orden natural en la ciudad las cuales, fuerzas y agrupamientos, deben ser aisladas y descritas por la ecología humana (Park y Burgess, 1967 [1925]). Bajo dichos supuestos, la segregación es entendida como expresión de interés alojada en las mismas clases o razas que optan por segregarse (Park y Burgess, 1967 [1925]), “*en la medida que sus requerimientos y modos de vida son incompatibles unos con otros (...) antagónicos entre sí*” (Wirth, 1962 [1938]: 28). De esto, aclara el mismo Wirth (1962 [1938]) que no se trata de procesos de atracción positiva entre población que se asemeja entre sí, sino como el resultado de diferencias de raza, ingresos o estatus económico, entre otros.

Su visión despolitizada de la espacialidad de la vida urbana, el desconocimiento de la geohistoria de las ciudades y la falta de una crítica explícita sobre la influencia del capitalismo post-industrial son las mayores críticas que encontró la ESCh (Soja, 2008). Aún con febles esfuerzos por relevar espacios intersticiales de la clase obrera en la vida urbana, como segunda derivada del crecimiento ecológico de la ciudad (Halbwachs, 2004 [1932]), no escapó este enfoque de las críticas antes mencionadas. Para los enfoques marxistas surgidos a partir de la crisis urbana en la década de 1960, estas falencias teóricas, sumadas a la naturalización de los procesos socio-espaciales urbanos en una suerte de metabolismo o evolución “darwiniana” de las ciudades (Park y Burgess, 1967 [1925]), eran imperdonables. Sin embargo, que el modelo de morfología socio-espacial desarrollado Burgess sea “*familiar a todos los estudiantes de sociología y geografía urbana*” en la actualidad (Soja, 2008: 137), habla del legado espacialista en los estudios urbanos y, en particular, en relación a la SR. No por nada se insiste en ocupar modelos de ese tipo tanto en antiguas reivindicaciones al ecologismo urbano (Dotson y Dotson, 1957), como para representar el presente de la ciudad fragmentada en Latinoamérica (Borsdorf, 2003).

En la segunda mitad del siglo XX, el interés que despertó en el marxismo el teorizar sobre la vida urbana y la SR se plasmó en trabajos de gran influencia para los estudios urbanos. En términos teóricos generales, la diferenciación socio-espacial

es vista por sus principales autores como la expresión espacial de los niveles de polarización social, producto de los procesos de neoliberalización en las ciudades y de acumulación del capital. Esta idea, alojada en los trabajos de autores como Lefebvre (1969, 2013 [1974]), Castells (1999 [1972]), Lojkine (1986 [1977]) y Harvey (1989), es desarrollada por cada uno siguiendo enfoques distintos –incluso opuestos–, según las diferencias político-ideológicas de sus programas.

En un primer orden de análisis, las diferencias dentro de sus planteamientos giran alrededor de la importancia que cada autor otorga al espacio y su relación con los procesos sociales. Lefebvre le atribuye mayor importancia al espacio, al resaltar la contradicción en su valor de uso y valor de cambio para la acumulación de capital (1969); vale decir, por su calidad de producto y medio de producción (2013 [1974]). Releva la SR como práctica de clasificación y distribución de las clases sociales –excepto la hegemónica– sobre el entorno urbano, “separándolas, prohibiendo los contactos entre ellas y sustituyendo los signos de su contacto” (2013 [1974], 407). Por el contrario, Castells (1999 [1972]) soslaya por completo la influencia del espacio dentro del conflicto de clases sociales. El análisis de SR lo realiza siguiendo la “distribución de los productos” y la “capacidad social de los sujetos”, definiéndola finalmente como la “expresión espacial” de la distancia social (1999 [1972]: 204).

Desde la perspectiva de la presencia de la espacialidad en las definiciones de SR, la tentación inmediata es situar la desarrollada por Lojkine (1986 [1977]) en la categoría de no-espacial, al relevar en la estratificación social los clivajes de clase (pp. 216-223). En cuanto al desarrollo teórico de Harvey (1989), aun cuando durante la década de 1970 se acercaba a una postura similar a la de Castells, tributa directamente de las teorías de Lefebvre. Harvey asigna un rol principal al espacio y la diferenciación residencial en la “fragmentación de la conciencia de clase” y en el reflejo de “las contradicciones en la sociedad capitalista” (1989: 118).

Sin embargo, las diferencias teóricas más importantes aparecen en la contraposición de los enfoques teóricos que utilizan para el análisis socio-espacial del entorno urbano. Lefebvre (1969), advirtiendo constantemente sobre el riesgo de

asignar una causalidad directa a distintos niveles de realidad social, propone en cambio que:

(...) el análisis de los fenómenos urbanos (...) exige el empleo de todos los instrumentos metodológicos: forma, función, estructura; niveles, dimensiones, texto, contexto, campo y conjunto, escritura y lectura, sistema, significativo y significado, lenguaje y metalenguaje, instituciones, etc. (p. 77).

Los tres primeros términos –forma, función y estructura– develan el enfoque de análisis que Lefebvre usa en su dialéctica, sin separar ni subordinarlos a priori entre ellos, evitando cualquier tipo de reduccionismo teórico (Lencioni, 2015). Esto lo diferencia de los planteamientos teóricos levantados por Castells, Lojkine y Harvey.

Castells (1999 [1974]) asume una postura estructuralista proveniente de un marxismo más ortodoxo, donde el espacio no tiene ninguna injerencia sobre los procesos sociales. Por su parte, Lojkine (1986 [1977]) y Harvey (1989) parten desde un enfoque dialéctico y relacional que va construyendo su teoría en torno a los procesos socio-espaciales –entre sujeto y objeto– sin la necesidad de definiciones preconcebidas. Lo curioso de las teorías urbanas de Lojkine y Harvey es que, a pesar de emular el proceder teórico de Marx –ambos lo explicitan–, sus conclusiones, en lo que a SR se refiere, distan bastante entre sí: Lojkine matiza el distanciamiento espacial de las clases sociales y pone en duda la reproducción del capital en el espacio, mientras Harvey ve en la especulación inmobiliaria procesos de acumulación con efectos claros en forma de SR y pérdida de la conciencia de clase.

Como se expone en el siguiente apartado de esta investigación, las posturas actuales que gobiernan el debate sobre el papel que juega la SR relación a dinámicas socio-espaciales de desintegración social, derivan directamente de la relevancia que los distintos autores otorgan a los enfoques recién descritos: es decir, a las influencias directas o indirectas de la morfología socio-espacial urbana, a las funciones socio-territoriales entre niveles locales y globales, y a las estructuras sociales y económicas que ciernen también en dichos niveles o escalas de análisis. Por consecuencia, destaco este análisis de enfoques teóricos sobre teoría urbana

y SR no sólo en su calidad de antecedente al estado del arte en la discusión, sino también por los aprendizajes particulares que se pueden obtener de este ejercicio. En tanto distintos enfoques de análisis teórico pueden llegar a conclusiones similares sobre la realidad socio-espacial o, al contrario, aproximaciones teóricas análogas pueden decantar en conclusiones divergentes, es deber de los investigadores urbanos –como en toda ciencia– cuidar la lectura que se realiza sobre los antecedentes empíricos.

2. Medición de segregación en Latinoamérica:

Del análisis anterior sobre los lineamientos históricos generales en torno a lo que es y significa la SR, y a fin de despejar las posibles contradicciones en su teorización y operacionalización en el marco de la realidad urbana latinoamericana, destacan los trabajos realizados por Rodríguez (2001) y Garrocho y Campos-Alanís (2013), a modo de resumir y consensuar las definiciones más utilizadas en la tradición investigativa sobre la diferenciación socio-espacial. En ambos aparece la definición de Massey y Denton (1988). Al introducir los conceptos de entorno urbano y grado de separación entre los distintos grupos sociales, Garrocho y Campos-Alanís relevan el carácter espacial intrínseco que debe existir a la hora de medir la segregación residencial.

Para el caso de la segregación residencial socioeconómica, una definición ampliamente citada en los estudios latinoamericanos sobre segregación es la que trabajan Sabatini *et al* (2001). Esta se entiende como el “*grado de proximidad espacial o de aglomeración territorial de las familias pertenecientes a un mismo grupo social, sea que éste se defina en términos étnicos, etarios, de preferencias religiosas o socioeconómicos*” (p. 27). Ateniéndose al análisis para distintos grupos socioeconómicos, para los autores existen tres dimensiones de la segregación residencial: (i) grado de concentración de grupos sociales en ciertas áreas de la ciudad; (ii) el grado de homogeneidad social de las áreas conformadas; y (iii) la percepción subjetiva de los residentes sobre las dos dimensiones anteriores (objetivas).

Sabatini y Sierralta (2006) defienden estas dimensiones, criticando la concepción tradicional de segregación residencial trabajada por Massey y Denton (1988), la cual reconoce cinco dimensiones objetivas: uniformidad, exposición, concentración, centralización y aglomeración –*clustering*–: mientras las dos primeras se asemejan a las dimensiones objetivas planteadas por Sabatini *et al* (2001), el resto no califica para los casos de segregación residencial para las ciudades latinoamericanas. La concentración definida como mayor densidad de población (en extremo hacinamiento), sería una dimensión de la pobreza, relacionada o no con la segregación; la cercanía al centro (inner-city) es una característica de las ciudades estadounidenses; y la aglomeración la entienden como sinónimo de segregación, pero referida a una escala específica de análisis (micro a meso).

Respecto a esto último, es pertinente resaltar que dichas apreciaciones alrededor de las dimensiones de SR planteadas por Sabatini son compartidas por Reardon y O'Sullivan (2004); adelanto esta observación, ya que utilizo dicho estudio como eje metodológico para la construcción de los índices de segregación a medir en esta investigación con los datos del GS de 2012.

3. Debate metodológico: separando la paja del trigo

Frente a los antecedentes teóricos recién expuestos, el debate sobre cómo medir la SR se instalaría –en principio– en relación a la capacidad de los índices sintéticos de capturar este fenómeno socio-espacial en las ciudades. Se le cuestiona a los índices de disimilaridad y aislamiento (Duncan y Duncan, 1955; Massey y Denton, 1988), en su versión más utilizada en las investigaciones sobre segregación residencial para Latinoamérica, de ser índices no-espaciales (Garrocho y Campos-Alanís, 2013; Ruiz-Tagle y López, 2014). Esta crítica, desde su óptica metodológica, dista de ser nueva. Los problemas operativos que trae medir la SR en índices sin componentes de interacción espacial, para cualquier escala de medición, nivel y geometría de agregación de datos, han sido ampliamente estudiados (Openshaw, 1983; White, 1983; Bailey y Gatrell, 1995; Lee y Culhane, 1998). Se abordan a continuación tales limitaciones metodológicas, para despejar cualquier duda

posterior sobre la relación entre índices de SR y las dimensiones tradicionales a relevar.

Se comprueba a través del problema del Tablero de Ajedrez la insensibilidad de los índices no-espaciales a la distribución espacial de las distintas unidades espaciales básicas de análisis (White, 1983). A su vez, al ser construidos estos índices a partir de relaciones determinísticas, la confiabilidad estadística de estos no es comprobable, dado que no es posible determinar los encuentros por proximidad espacial a modo de probabilidad. Incluso los datos censales por manzana, al promediar variables sociales y económicas, merecen un análisis de este tipo. Por último, el problema de la unidad espacial modificable –MAUP, por su sigla del inglés–, genera un sesgo de difícil solución (Openshaw, 1983). Al incorporar variaciones al valor que otorgan los índices de SR según la forma y nivel de agregación de las unidades espaciales básicas de análisis, introduce el cuestionamiento válido sobre cuál es la escala propicia para analizar la diferenciación socio-espacial. Dicho esto, ¿quedan completamente invalidados los estudios sobre SR que analizan cuantitativamente el fenómeno a través de los índices de disimilaridad y aislamiento?

Un estudio que revela de mejor manera las dimensiones que relevan los distintos índices, espaciales y no espaciales, es el que realizan Lee y Culhane (1998), donde se simulan situaciones teóricas y reales. Los casos reales se miden en relación a los niveles de segregación que arrojan los distintos índices, respecto a la población negra, población pobre y población indigente (*homeless*, sin hogar) para distintos distritos de Nueva York. Fuera de la presentación de resultados, son de especial interés las conclusiones relativas a las diferencias entre las distintas propuestas metodológicas: Por una parte, el índice de agrupamiento (*clustering index*) propuesto por los autores, confiere relevancia a las distribuciones de agrupamiento del tipo enclave, el cual se relaciona mejor con efectos de guetización (contigüidad de cierta aglomeración y separación de esta con otros grupos sociales), mientras que el índice de autocorrelación global de Moran y el de proximidad espacial de

White, dan mayor prioridad a la contigüidad en cuanto al nivel de vecindad entre las distintas unidades espaciales.

Lee y Culhane reconocen que los tres índices de autocorrelación espacial tienden a subestimar el efecto de concentración de un grupo en un mismo espacio, situación que es capturada de mejor manera por el índice de disimilaridad. Esto obliga, implícitamente, a una segunda lectura sobre el grado de relevancia que un índice de segregación residencial otorga a la contigüidad (vecindad) y proximidad (distancia) entre las unidades básicas de análisis, y su impacto sobre la importancia que le resta a su composición social interna. Con tales resultados, ¿se puede determinar “qué sucede” con la segregación residencial?

Desde la perspectiva del MAUP, Openshaw (1983) determina que un reagrupamiento de las unidades espaciales modificará los niveles de segregación. En su análisis, Openshaw revisa que este problema tiene dos dimensiones: la escala y la forma (geometría) de agregación de datos. Este problema, planteado como un dilema para los índices no espaciales, también lo es padecido por los índices de autocorrelación espacial (Wong, 2004). Garrocho y Campos-Alanís (2013) argumentan sobre formas de reducir este problema para los índices de autocorrelación (global de Moran y LISA, específicamente), pero también puede ser abordado desde los índices no-espaciales (Openshaw, 1983; Sabatini *et al*, 2001; Sabatini y Sierralta, 2006; Rodríguez, 2013).

Al proponer un procedimiento iterativo para agregar unidades espaciales en nuevas zonas de análisis para reducir el MAUP, Openshaw (1983) intuye la necesidad teórica de determinar previamente los fenómenos que se quieren predecir, y que estos, por su especificidad, implicarían distintas escalas y formas de zonificación. Asimismo, Sabatini *et al* (2010) y Rodríguez (2001) resaltan la importancia de la escala de análisis, dependiendo del fenómeno socio-territorial ligado a la segregación residencial.

CAPÍTULO III.
METODOLOGÍA DE LA INVESTIGACIÓN
CLUSTERIZACIÓN JERÁRQUICA PARA ÍNDICES LOCALES DE
SEGREGACIÓN RESIDENCIAL

Se presenta a continuación la metodología para estudiar las distintas escalas que toma el fenómeno de SR según la dimensión en particular que se quiere medir y el índice a utilizar con ese fin. Sin embargo, para efectos de esta investigación, sólo se presentarán los resultados preliminares producto del diseño base de esta misma. Del análisis de dichos resultados nacen las directrices a seguir para completar una rutina de regionalización capaz de encontrar una escala óptima de análisis.

1. Índices espaciales de SR: capturando sus principales dimensiones en el entorno local

Acorde a la discusión metodológica en torno a las dimensiones que deben representar los índices de SR, y tal como fue adelantado en la introducción de esta tesis, planteo adaptar y medir tres de los cinco índices espaciales que Reardon y O'Sullivan (2004) desarrollan en su trabajo: una medida de exposición –aislamiento, xPx_r – y dos medidas de uniformidad –disimilaridad, D_{xr} , y el índice de la teoría de información de Theil, H_r –.

Desde el punto de vista de la producción investigativa en SR, los índices no espaciales de disimilaridad y aislamiento han sido ampliamente estudiados; bien asimilados por tratarse de índices fáciles de usar e interpretar (Sabatini y Sierralta, 2006). La idoneidad de utilizar estos índices y extender en su desarrollo aparece, en primera instancia, gracias a su comparabilidad con mediciones anteriores de segregación en el GS. No obstante, en el último decenio es el índice de la teoría de información el que cuenta con mejor reputación (Lee *et al*, 2008; Monkkonen y Zhang, 2014; Östh, Malmberg y Andersson, 2014; Fowler, 2015), lo que se condice con ser la “mejor medida de uniformidad espacial” de los índices propuestos y evaluados por Reardon y O'Sullivan (2004: 153).

Propuesta de índices espaciales locales

Similar a como operan Reardon y O'Sullivan (2004), para una región R se define $r_p \subseteq R$ el entorno local de la unidad espacial p tal que $p \in R$. A su vez, la pertenencia de otra unidad espacial $q \in R$ al entorno r_p se define a través de una función de vecindad $\varphi(p, q)$ no negativa tal que $\varphi(p, q) = \varphi(q, p) \wedge \varphi(p, p) = \varphi(q, q) \forall p, q \in R$. En términos simples, esto quiere decir que si q es parte del entorno local de p , también es cierto que p es parte del entorno local de q .

El nivel de disimilaridad del grupo x en su entorno local r se calcula mediante la expresión:

$$D_{xr} = \frac{1}{2} \sum_{i=1}^n \left| \frac{t_{xi}}{T_{xr}} - \frac{t_i - t_{xi}}{T_r - T_{xr}} \right| \quad \forall i \in r$$

Donde

t_i = población total en unidad espacial i

t_{xi} = población del grupo x en unidad espacial i

T_r = población total en entorno local r ($\sum_i t_i$)

T_{xr} = población total del grupo x en entorno local r ($\sum_i t_{xi}$)

Esta definición es idéntica a la planteada por Massey y Denton (1988) para medir disimilaridad con sólo 2 grupos, con la única diferencia de que la función de proximidad permite tanto variabilidad en la definición del entorno local según el objeto de investigación, como la pertenencia de una misma unidad espacial a distintos entornos locales. Por otro lado, la diferencia más importante con el índice propuesto por Reardon y O'Sullivan (2004) –fuera de ser una versión desagregada para cada entorno local– es que mide el valor total objetivo de disimilaridad como proporción de su máximo teórico en vez de calcular su valor relativo a la uniformidad de la región R bajo estudio. La diferencia teórica de estos supuestos es importante: para Reardon y O'Sullivan (2004), su índice de disimilaridad mide “cuan diferente, en promedio, es la composición de los individuos en su entorno local respecto a la composición de la población como un todo” (p.141); mi propuesta, en cambio,

recupera la interpretación original del índice, al representar la “proporción de individuos que debiesen moverse para lograr una completa integración” (pp.140-141).

El cálculo de el índice de aislamiento del grupo x en su entorno local r se define de la siguiente forma:

$$xP^*x_r = \sum_{i=1}^n \left[\frac{t_{xi}}{T_{xr}} \right] \left[\frac{t_{xi}}{t_i} \right] = \sum_{i=1}^n \left[\frac{t_{xi}}{T_{xr}} \right] \pi_{xi} \quad \forall i \in r$$

Donde

$$\pi_{xi} = \text{proporción del grupo } x \text{ en unidad espacial } i \left(\frac{t_{xi}}{t_i} \right)$$

En este caso, la formulación del índice –siguiendo las modificaciones estudiadas por Lieberson y Carter (1982) al modelo de Bell (1954)– es una adaptación de escala respecto a la extensión espacial que hacen Reardon y O’Sullivan (2004); mantiene la proporción del grupo estudiado dentro de cada unidad espacial del entorno para el cálculo de probabilidad de interacción, mientras que deja al arbitrio de la función de vecindad el cómputo de las unidades consideradas y la población total. Ciertamente, esta adaptación es más similar al planteamiento no espacial y pierde el uso de proporción de grupo en el entorno, pero eso es una limitante difícil de eludir como índice local en vez de una ponderación para toda la región de estudio. Soluciones multiescales son posibles –como definir entornos anidados, por ejemplo– pero debería resolverse primero, tal como considera el foco de esta tesis, el encontrar métodos efectivos para definir una escala óptima de análisis.

Por su parte, el índice de la teoría de información de Theil (1972) se adapta de manera especial para mantener su interpretabilidad a nivel local. En primer lugar, definimos dos funciones de entropía multigrupal; una para el entorno local r_p y otra general para la región R :

$$E_{r_p} = - \sum_{m=1}^M (\pi_{r_p m}) \log_M (\pi_{r_p m})$$

$$E_R = - \sum_{m=1}^M (\pi_m) \log_M (\pi_m)$$

Donde

$\pi_{r_p m} = \text{proporción del grupo } m \text{ en el entorno } r_p$

$\pi_m = \text{proporción del grupo } m \text{ en el en la región } R$

Ambas fórmulas miden la diversidad por composición de los grupos presentes en sus respectivas escalas. Tomarán un valor máximo de 1 cuando la presencia de cada grupo sea idéntica, y 0 con la presencia de un solo grupo. Luego, el índice de la teoría de información se calcula como la diferencia entre el valor máximo teórico de 1 –completa segregación– y la razón ponderada de las entropías por su población efectiva:

$$H_r = 1 - \frac{1}{T_r E} \sum_{i=1}^n t_i E_{r_i}$$

Difiere al índice planteado por Reardon y O'Sullivan (2004), en el sentido que se modifican las poblaciones que ponderan la razón entre entropías. La idea tras esta forma de adaptar el índice a nivel local sigue la lógica de que la entropía representa la experiencia de diversidad de la población en cada unidad espacial i sobre su entorno r ; hace sentido entonces que sea ponderado por la población de cada unidad espacial t_i y no por la población total T_r de su entorno local. La base de comparación es la diversidad teórica ideal que debiese experimentar en conjunto todo el entorno r , bajo el supuesto de que distribuye siguiendo la composición general de la región R .

La justificación se piensa para determinar la distancia para lograr la mejor distribución posible para la región en su conjunto. Por lo mismo, este índice permite valores negativos cuando el entorno r se compone con mayor uniformidad que la región estudiada, asigna el valor 0 cuando el entorno local se compone de manera idéntica a la región R , y el valor 1 cuando sólo hay uno de todos los grupos presente en dicho entorno.

Al incorporar este índice, de carácter multigrupal y relativo a la composición total de la región, suma la investigación un elemento relevante a la hora de evaluar la escala

de los procesos socio-territoriales de SR en relación a la concentración de los estratos socio-materiales medidos con el ISMT, sin el sesgo de selección a la hora de elegir un corte sobre como dividir a los estratos de la población en dos grupos de estudio.

Entorno local y función de proximidad

La función de proximidad que define los límites del “entorno local” se torna un punto clave para los índices propuestos. De esta se desprende la sensibilidad de los resultados de la medición.

A partir de la necesidad de encontrar la escala que mejor refleje la espacialidad de la dimensión –uniformidad o exposición– de SR bajo estudio, se precisa un umbral de 500 metros para delimitar la vecindad del entorno local que mide los índices locales de segregación para cada manzana. El supuesto involucrado en esta decisión –distancia estándar en estudios de planificación de movilidad peatonal– es asegurar una medición dentro de la escala barrial, entendida como el entorno más inmediato o, en palabras de Saraví (2004), “el primer encuentro público al abrirse la puerta de lo privado” (p.35). En ese aspecto, las diferencias de vecindad por distancia de en los primeros 500 metros son mínimas. Por ende, se asigna una relación de proximidad uniforme a todas las manzanas que caen dentro de dicha definición de entorno local.

Lee *et al* (2008), por ejemplo, miden segregación racial con el índice de la teoría de información para 100 áreas metropolitanas utilizando una función de distancia inversa. Esta otorga mayor peso a las vecindades más cercanas a través de un kernel bidimensional que aproxima esta relación a una curva normal dentro de un radio establecido. En otra investigación, Östh, Clark y Malmberg (2015) miden el índice de aislamiento para Los Ángeles y tres áreas metropolitanas de Suecia definiendo el entorno local según barrios con cantidad de vecinos constante – método de *k-nearest neighbor*–, variando los pesos espaciales según la densidad de población en cada unidad espacial. Si bien ambas pesquisas se aproximan de

forma exhaustiva al concepto de entorno local, las dos pertenecen al grupo de publicaciones que miden SR desde una perspectiva multiescalar que compara el resultado de medir segregación a distintas escalas con un mismo índice. ¿Es válido hacer dicha comparación?

Sin ir más lejos, Sabatini *et al* (2010) hacen el mismo tipo de análisis multiescalar para el GS en el periodo intercensal 1992-2002. Utilizando los índices no espaciales de disimilaridad y aislamiento, llegan a similares conclusiones con las investigaciones recién citadas en cuanto “se trata de un efecto esperable del cambio de escala antes que una distorsión o imperfección del índice” (p.25). Por el contrario, sostengo que esta afirmación tiene dos falencias importantes de relevar.

Desde el punto de vista teórico, dicha perspectiva es contraria a considerar la escala geográfica de análisis como categoría epistemológica. Como ya se ha discutido, los índices de SR bajo estudio se definen acorde a dimensiones socio-espaciales objetivas particulares que cuentan con su propia espacialidad. Luego, pierden su sentido los índices de SR si se aplican estos en escalas que no se corresponden con su espacialidad.

En cuanto al ámbito práctico/metodológico, esa postura es imprecisa. Tanto para mediciones realizadas con índices espaciales y no espaciales, los cambios por escala de agregación de los datos no son comparables por efectos del MAUP; o visto de otro modo, ante la incapacidad de cuantificar el efecto del MAUP sobre los resultados de cambio de escala, no hay evidencia para sostener que existe un efecto propio del cambio de escala.

Es imprescindible cautelar el objeto de estudio entonces, se defina como segregación étnica, socio-económica, educacional u otra; lo mismo con su escala de análisis, acorde a los procesos que existen tras de esta. En ese sentido, al utilizar una herramienta metodológica como es la definición de entorno local propuesta por Reardon y O’Sullivan (2004) –la cual logra sortear, en gran parte, el problema del MAUP– no deben descuidarse los preceptos teóricos que la sostienen.

Por consiguiente, adelanto una aclaración para evitar malos entendidos: pareciera que la misma definición de entorno local queda en entredicho, pero no es el caso. El producto de esta investigación considera el sesgo en la definición de entorno local, entendiéndolo que sus resultados entran en la categoría de exploración de los datos, quedando a disposición para su interpretación más que cualquier análisis de tipo inferencial. Por otro lado, traigo a colación uno de los objetivos de esta tesis: plantear, acorde a sus resultados, directrices metodológicas para encontrar escalas óptimas de análisis para los estudios de SR; en ese contexto, esta investigación es, antes que nada, insumo para encontrar nuevos métodos de análisis de la SR basados en algoritmos de aprendizaje no supervisado.

2. Evaluando la escala mediante un algoritmo de regionalización jerárquica

Como caso particular de clusterización espacial, los algoritmos de regionalización derivan directamente de procedimientos generales de optimización utilizados para identificar y agrupar datos similares, garantizando que las observaciones en diferentes grupos difieran lo máximo posible entre ellas. (James, Witten, Hastie y Tibshirani, 2013; Garretón y Sánchez, 2016).

La forma más común de medir la diferencia entre observaciones al interior de un *cluster*—la cual se utiliza en el algoritmo escogido en esta investigación— es a través de una distancia Euclidiana cuadrática. En notación de James *et al* (2013), se operacionaliza

$$W(C_k) = \frac{1}{|C_k|} \sum_{i,i' \in C_k} \sum_{j=1}^p (x_{ij} - x_{i'j})^2$$

Donde $W(C_k)$ es la magnitud de esa diferencia entre observaciones que pertenecen al *cluster* C_k ; $|C_k|$ indica el número de observaciones en el k -ésimo *cluster*, mientras que x_{ij} y $x_{i'j}$ son el par de observaciones i e i' de la variable j ; esto se simplifica al tratar cada variable —índice de segregación— por separado. Luego, se busca optimizar la expresión

$$\min_{C_1, \dots, C_k} \left\{ \sum_{k=1}^K \frac{1}{|C_k|} \sum_{i, i' \in C_k} (x_i - x_{i'})^2 \right\}$$

Para un algoritmo de regionalización jerárquica que utiliza una rutina *bottom-up*, se resuelve con un paso inicial en que considera cada observación como su propio *cluster* y, bajo una restricción de vecindad espacial dada, busca las dos observaciones con menor distancia en la variable evaluada. Luego, ese par de observaciones las agrupa en un único *cluster*, para repetir iterativamente el proceso de evaluación y fusión hasta agrupar todas las observaciones en dos regiones.

A diferencia de las observaciones individuales, para medir la distancia entre los *cluster* se debe elegir un método *linkage* que los vincule. Los principales métodos se resumen en la tabla 3.1.

Método de vinculación	Descripción
Complete	Asigna la máxima distancia posible entre todos los pares de observaciones individuales inter- <i>cluster</i> , en cuanto a las variables medidas.
Single	Inverso al <i>complete linkage</i> , asigna la distancia entre el par de observaciones individuales inter- <i>cluster</i> más similares entre sí.
Average	Es el promedio de las distancias entre todos los pares posibles de observaciones inter- <i>cluster</i> .
Centroid	Para cada iteración, computa la media de las observaciones dentro de cada <i>cluster</i> y para medir directamente la distancia
Ward's method	En vez de calcular la distancia entre distintos <i>cluster</i> , computa todas sus combinaciones posibles con sus respectivas sumas cuadráticas internas, para escoger la combinación que entregue la menor suma.

Tabla 3.1. Métodos de vinculación comunes en regionalización jerárquica
Fuente: elaboración propia a partir de James *et al* (2013) y Anselin (2018)

El método debe ser coherente con su fin; se pretende evaluar la escala de la SR a partir de identificar y unir aquellas manzanas que experimentan un mismo nivel de segregación. Además, al ser medida la SR a nivel de entorno, es esperable –pero no la regla– que dos manzanas vecinas tengan valores similares por compartir buena parte de ese entorno. No importaría entonces que dos manzanas similares quedaran en *cluster* separados, si esto refleja un cambio paulatino del nivel de SR;

por el contrario, manzanas disímiles en su nivel de segregación no pueden quedar agrupadas. Luego, el método de Ward parece más adecuado, conforme minimiza la suma de los cuadrados al interior de cada grupo de manzanas similares, sin prestar mayor atención a la distancia entre distintos *cluster*.

Dentro de las ventajas de la regionalización jerárquica que mencionan James *et al* (2013) respecto de un método de *k-means*, es que no se necesita entregarle un número predeterminado de *clusters* para operar, además de entregar una “atractiva representación de las observaciones” mediante un dendograma (p.390). En cuanto a las posibles soluciones subóptimas por agrupación anidada de los *clusters* (p.394), representa un problema genérico del algoritmo de clusterización cuando no se trata con datos espaciales restringidos a una función de vecindad.

Las limitantes del método radican más en la variable de segregación evaluada: al asignar todas las unidades a algún grupo, podría generarse zonas con *clusters* espurios ante valores de escasa segregación; también podría resultar en agrupaciones indebidas cuando la escala de SR es igual o menor a la escala del entorno local definido para calcular la segregación en el paso anterior.

CAPÍTULO IV.
ESPACIO Y ESCALAS DE LA SEGREGACIÓN
RESIDENCIAL EN SANTIAGO
ANÁLISIS Y DISCUSIÓN DE LOS PRINCIPALES RESULTADOS DE
LA INVESTIGACIÓN

Para la obtención de estos resultados, se utilizó un corte de estratificación que incorpora en el grupo estudiado – índices de disimilaridad y aislamiento – a los estratos socio-materiales E, D y C3 del ISMT. La elección se sustenta, primero, en torno a la interpretabilidad que se puede obtener mediante el indicador escogido para la estratificación; el no contar con una variable de estratificación continua y de mayor confianza estadística para predecir los ingresos, los resultados de SR para las clases más vulnerables podrían verse sesgados. En segundo lugar, este corte coincide con uno de los utilizados por Agostini *et al* (2016), del 64% más pobre, el cual estaba exento del pago de contribuciones a mediados de los años noventa (p.168), dato importante para efectos de la distribución socio-espacial de los distintos estratos por su especificidad territorial.

1. El patrón de segregación residencial en el Gran Santiago

El panorama general que exhibe el Santiago de 2012 en cuanto a SR concierne, en línea con lo sostenido por Sabatini *et al* (2010) y Sabatini, Cáceres y Rasse (2013), muestra claros indicios de transformación respecto al patrón tradicional de las últimas décadas, caracterizado por grandes zonas homogéneas de pobreza y la concentración espacial de los estratos socioeconómicos de más altos ingresos en el conocido “cono de alta renta” (Ducci, 2000; Sabatini *et al*, 2001; Hidalgo, 2004; Sabatini *et al*, 2010). Sin embargo, sostengo a partir de las mediciones realizadas que, dicha transformación, tiene matices y bemoles que deben explicitarse.

Proximidad cromática

Sostienen Sabatini *et al* (2013) el importante rol que ha jugado la concentración del capital inmobiliario en los cambios del patrón de segregación, desarrollando proyectos destinados a segmentos altos y medios en “sectores tradicionalmente asociados a otros grupos sociales, en la medida que la escala de los proyectos cambia por completo las características del territorio” (p.223). Este proceso, que suele circunscribirse en contextos urbanos de periferia popular por lo relevante en

términos de integración o exclusión social respecto a los sectores más vulnerables de la población (Rasse, 2015; Ruiz-Tagle, 2016), tiene en realidad su correlato en casi toda el área metropolitana de Santiago.

Exceptuando el cono de alta renta y sectores en la periferia oeste y sur de Santiago –sectores de Cerro Navia y Renca al poniente; La Pintana, Bajos de Mena en Puente Alto y sectores en San Bernardo y El Bosque–, la diversidad en la composición socio-material a nivel local trasciende a dicha microescala. La figura 4.1. da cuenta de aquello a través del índice multigrupal de Theil; al ser una ponderación de nivel de entropía, medida esta en escala logarítmica con los cinco estratos socio-materiales del ISMT, parece difícil de interpretar en primera instancia. Para despejar dudas, la tabla 4.1. muestra algunos valores de referencia:

<i>Grupos presentes</i>	<i>Proporción</i>	<i>H_r (aproximado)</i>
4 de 5	1:1:1:1	0,12
3 de 5	1:1:1	0,30
3 de 5	2:2:1	0,33
2 de 5	1:1	0,55

Tabla 4.1. Valores de referencia del índice de Theil (5 grupos).
Fuente: elaboración propia

**Segregación Residencial local en
Área Metropolitana de Santiago, 2012
Índice - T. de Información de Theil (Hr)**

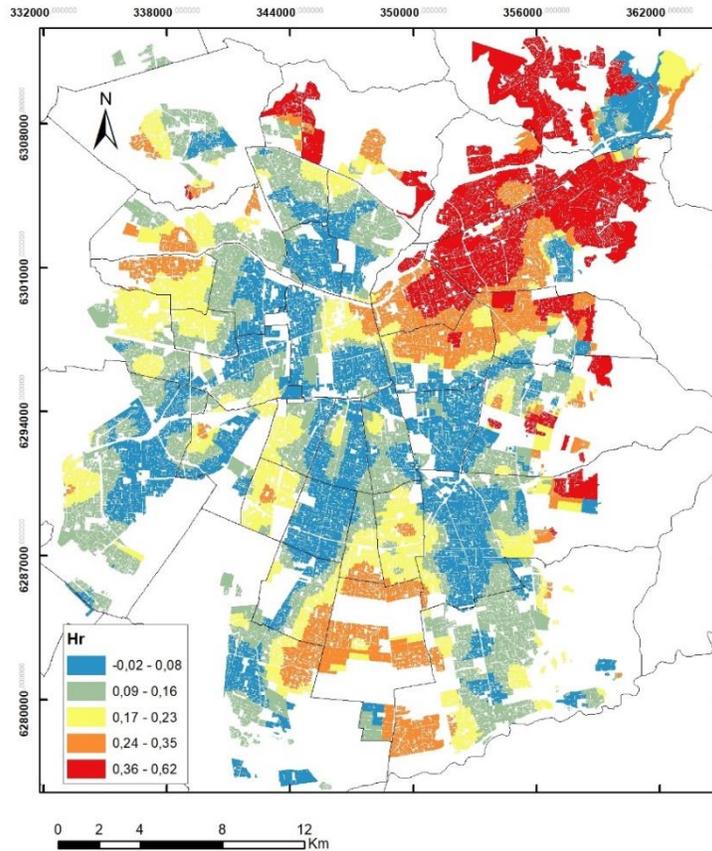


Figura 4.1.

Mapa del índice de Theil a escala barrial. Santiago, por manzana, 2012.
Fuente: elaboración propia a partir del ISMT (OCUC, 2018)

Empero esta fotografía de expansión de la diversidad socio-material, no se puede inferir en ningún caso que dicha dispersión y mezcla social signifique una reducción de escala de la segregación –recordar que la segregación tiene más de una dimensión–. Sí tal vez, su intensidad en cuanto a la concentración espacial de la clase dominante; pero por cómo se mide, comparando la entropía a escala barrial con la composición total del GS, esconde este índice las condiciones particulares de tal fenómeno en torno a la calidad de esa uniformidad socio-espacial. Con eso en mente, relevo los resultados que nacen de aplicar el índice de disimilaridad local

como complemento insoslayable para el análisis de la concentración espacial de las distintas clases socio-materiales, resumido en la figura 4.2.

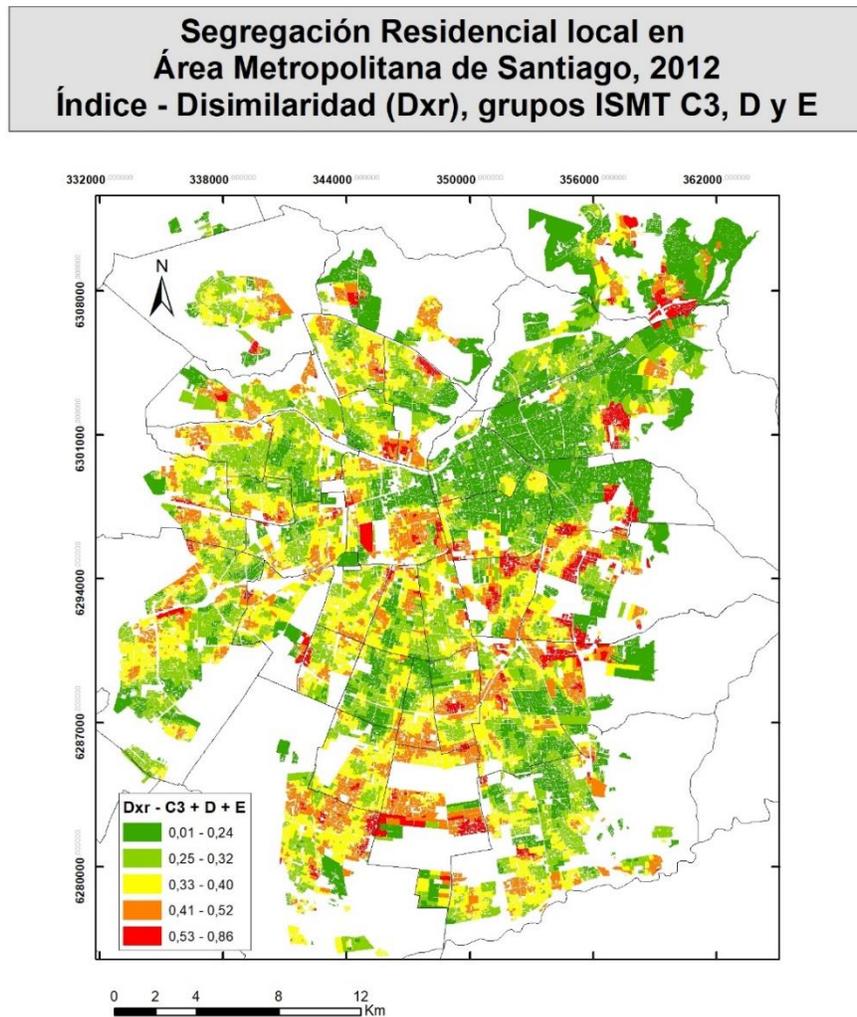


Figura 4.2.

Mapa de disimilaridad a escala barrial. Santiago, por manzana, 2012

Fuente: elaboración propia a partir del ISMT (OCUC, 2018)

Con ambos mapas se aprecia que, en la configuración socio-espacial del área metropolitana de Santiago, existen distintos tipos y formas de proximidad entre las distintas clases sociales, variando tanto en intensidad como en su escala geográfica. Antes de describir estas proximidades, cabe recordar que el índice de disimilaridad local debe interpretarse en su calidad de indicador de concentración relativo a los estratos sociales presentes en la microescala que está medido.

El ejemplo más claro se da en gran parte del cono de alta renta, donde no existe esa proximidad; de la figura 4.1 se extrae, sin dobles lecturas, que apenas se percibe la presencia de un tercer grupo no contribuyente en el área con índice de Theil $H_r > 0,36$. Lo que si se puede afirmar dentro del cono de alta renta es que, en los sectores que logran penetrar las clases medias, estas no lo hacen de forma concentrada en áreas específicas, sino que se mezclan de forma proporcional en el espacio; pareciera ser que, como proceso de transformación, la llegada de las familias de menores ingresos en barrios que nacieron como exclusivos o “de élite” (Sabatini *et al*, 2010: 31) ocurre de forma acotada. Lo cierto es que sólo con los datos cuantitativos de un periodo censal en particular, no se puede inferir exactamente lo que sucede en la frontera del cono de altos ingresos, donde es muy posible que coexistan y traslapen procesos de penetración de capas medias con una expansión un poco menos concentrada de las clases dominantes hacia los sectores más próximos de su residencia histórica de las últimas décadas.

Entre las variopintas geografías de proximidad, en la comuna de Santiago, por ejemplo, el sector de mayor heterogeneidad socio-material, alrededor de avenida Matta y hacia el sur, un nivel de disimilaridad local media-alta es muestra de producción inmobiliaria destinada a estratos más altos, lo que se replica en mayor intensidad y menor escala al sur de la frontera entre Recoleta e Independencia, alrededor de La Vega. Otros dos casos registrados, con resultados similares a nivel de estos dos índices, pero distintos a los ya mencionados – y entre sí –, se ven en La Florida (Ruiz-Tagle, 2016) y Macul (Atisba Monitor, 2017).

Por un lado, en el área de La Loma de La Florida, conviven vivienda social junto a condominios cerrados. Las mediciones de diversidad barrial realizadas por Ruiz-Tagle (2016: 103-104) coinciden con lo que muestra el índice de Theil, mientras que un índice de disimilaridad local medio-alto levanta el carácter de habitar en forma concentrada de los grupos en el entorno urbano.

En Macul, tomado de un registro informativo realizado para el diario La Tercera (Atisba Monitor, 2017), se perciben dos villas de largo trayecto histórico, Santa Julia y 23 de Enero, insertas en un entorno urbano donde residen estratos medios y altos.

Aquí, los altos índices de disimilaridad local representan la concentración en la microescala de los grupos populares.

Como ejemplo final, en contraste a estas “proximidades distantes” –frase que bien podría perder su calidad de oxímoron en el GS- emergen algunos patrones de coincidencia entre valores altos y medios de diversidad, medida por el índice de Theil, y baja disimilaridad. Probablemente estos sean los casos más difíciles de interpretar.

Uno de esos casos es la comuna de La Cisterna. Sus altos niveles de diversidad, con presencia de todos los estratos socio-materiales y su baja disimilaridad local, podría ser producto de movilidad social endógena en el sector o la llegada de nuevos vecinos de clase media y alta en conjuntos socialmente diversos. O ambas situaciones. En contexto, al comparar con San Miguel, se ve que alrededor de la Gran Avenida, específicamente en el sector de Lo Vial, hay un proceso extendido de gentrificación que expulsa a los estratos medios y bajos, casi por sobre los límites barriales de entorno local, según fueron definidos para realizar las mediciones. Si no, no se explica tal disminución puntual de la diversidad, junto con una disminución equivalente en la disimilaridad local –concentración homogenizante–. Luego, la alta diversidad a lo largo del eje de la Gran Avenida podría evidenciar los primeros pasos de un proceso de gentrificación de mayor escala.

Ante tales situaciones, determinar que sucede en este tipo de situaciones sin un trabajo de investigación de campo adecuado, no pasa de la especulación con datos empíricos pivotados con los imaginarios subjetivos que se tiene de la ciudad. Recordando los límites que tiene esta investigación, también podría tratarse de algunos sesgos de selección en la construcción de la variable de estratificación. Por gran herramienta que parezcan ser estas mediciones, no hay que perder de vista que es sólo una parte de la realidad socio-territorial.

Interacciones diatónicas

Completa el paisaje de la SR en el Santiago de 2012 las interacciones que experimentan ambos grupos socio-materiales entre si. Aun cuando la figura 4.3. muestra un patrón similar al de la figura 4.1., su interpretación no es, bajo ningún caso, la misma.

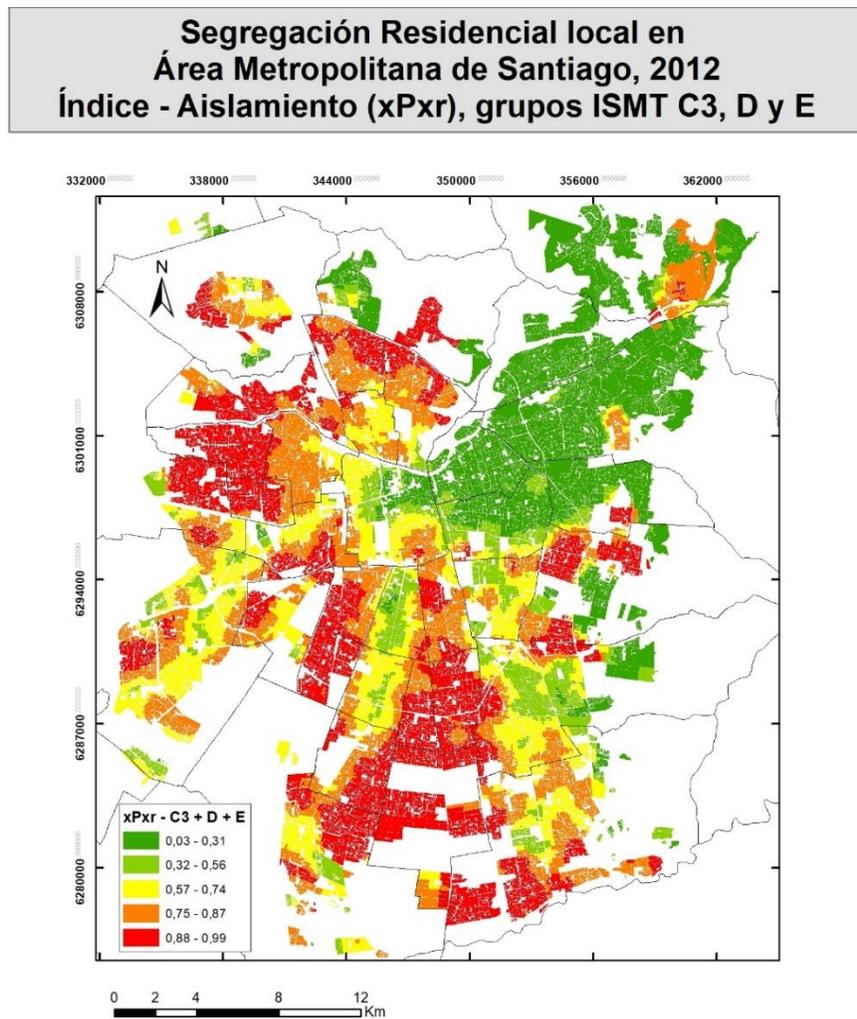


Figura 4.3.

Mapa de aislamiento a escala barrial. Santiago, por manzana, 2012

Fuente: elaboración propia a partir del ISMT (OCUC, 2018)

La correcta lectura de este mapa viene, en primera instancia, de entender que nuestro grupo de referencia corresponde a alrededor del 63% más pobre de la población del GS. Luego, los valores extremos que toma el aislamiento local –primer y quinto intervalo en la figura 4.3.– nada nos dice que no sepamos de la figura 4.1.;

los sectores con más altos niveles de índice de Theil arrojarán valores extremos en el índice de aislamiento, según el grupo predominante en aquel espacio de poca diversidad.

El valor agregado del índice de aislamiento local al análisis realizado en la sección anterior se da en las fronteras de ese aislamiento –o exposición– extremo. Dado que se calcula ponderando al grupo medido respecto sólo del entorno, se interpreta directamente su resultado como la proporción promedio a escala barrial.

Así, un aislamiento medio-alto o medio-bajo indica mayor o menor proporción, respectivamente, del grupo medido en dichas interfases de transición; por su parte, el índice de Theil dirá si dicha mezcla mantiene la proporción de los cinco estratos socio-materiales –como en Quinta Normal y La Cisterna, por ejemplo– o si esa interacción espacial sucede a por medio de expulsión socio-espacial –sector Lo Vial en San Miguel–. Por último, el índice local de disimilaridad dirimirá sobre si los niveles de interacción y diversidad espacial suceden dispersos en el entorno –conjuntos mixtos, por ejemplo– o bien son parte de transformaciones inmobiliarias dirigidas por capacidad de pago/endeudamiento.

Primera discusión: índices y dimensiones de SR

Los resultados preliminares de esta investigación son muy satisfactorios; mientras los índices locales medidos mantienen el significado original para el que fueron desarrollados, logran relevar los distintos procesos de SR que ocurren simultáneamente en la región bajo estudio. A su vez, diferencian claramente las dimensiones que pretenden medir –exposición y concentración del grupo seleccionado– sin perder el carácter más importante a la hora de medir segregación: su espacialidad.

Pero debo señalar, como explicité en el desarrollo de esta tesis, que al nivel que están desarrollados estos índices –esto es, como índices espacializados a un entorno local arbitrario– no pasan de ser buenas herramientas de exploración de los patrones de SR, complementarias entre ellas y susceptibles de segundas lecturas,

mientras no se interpreten acorde a antecedentes más robustos sobre los procesos socio-espaciales de SR. Cada uno de estos tres índices debe revisarse con cuidado.

El índice local de disimilaridad, como su original (Duncan y Duncan, 1955; Massey y Denton, 1988), representa la proporción de población que debiese mover su residencia en ese entorno para lograr integración completa. Al estar definido a escala barrial, no es posible sumar, a modo de puzle, su efecto entre unidades contiguas para hablar de un fenómeno a mayor escala. Tampoco parece opción medirlo para unidades geográficas más extensas, como zona censal, porque dentro de dicha agregación de datos se perdería el sentido original de concentración y dispersión bajo el manto de las proporciones en escalas que traspasan el barrio.

Por el contrario, el índice de aislamiento (Bell, 1954; Lieberman y Carter, 1982) muestra mejor adaptabilidad sin perder su significado como intensidad –o probabilidad– de la interacción de los grupos sociales en el espacio. En ese sentido, se intuyen oportunidades al medirse con distintos cortes de estratificación social: de contar con un método confiable para obtener escalas óptimas de análisis, no sería raro que el nivel de agregación de los datos espaciales siguiera esa escala para medir procesos en la dimensión de exposición socio-espacial.

Finalmente, los resultados obtenidos con el índice de teoría de la información de Theil (1972) merecen precisar su relación con la dimensión de uniformidad o concentración espacial, bajo la cual está catalogado. No es casualidad, a mi entender, que el patrón geográfico que releva para el GS este fuertemente correlacionado con el que muestra el índice de aislamiento. ¿Puede entenderse la diversidad socio-espacial como medida de uniformidad y de exposición? En su investigación sobre la extensión de índices sintéticos de SR a mediciones multigrupales, Reardon y Firebaugh (2002) realizan un ejercicio operativo donde derivan de distintas fórmulas genéricas los índices de segregación. Para el caso del índice de Theil, este puede derivarse de medidas de desproporcionalidad –tal cómo el índice de disimilaridad–, de asociación, y de diversidad –como el índice aislamiento, en su versión para dos grupos–. Fuera de sólo ser otra propiedad

operativa, este hecho demuestra que el índice de Theil mide ambas dimensiones, uniformidad y exposición.

A modo de hipótesis sobre el resultado que emerge del análisis de la SR en el GS de 2012, sostengo que el índice local de Theil que he desarrollado revelará patrones de SR de acuerdo a la interacción de ambas dimensiones objetivas. Pero, por otra parte, sostengo también que no es posible valerse sólo de este índice para medir SR; debe evaluarse a la par de los índices de disimilaridad y aislamiento, los cuales si son capaces de medir cada una de las dimensiones objetivas de SR por su propio mérito.

2. Evaluando las escalas de la SR

Regionalización jerárquica en tres comunas del GS

Debido a la complejidad en lo que respecta a este tipo de rutinas en cuanto a capacidad de cómputo, el algoritmo de regionalización jerárquica descrito en el capítulo metodológico de esta tesis se ejecutó para tres comunas distintas del GS de 2012. Para escogerlas, se midió para cada comuna su índice global de autocorrelación I de Moran según el puntaje continuo y normalizado del ISMT, para encontrar tres tipologías distintas de comuna: una con un alto índice de autocorrelación – $I \approx 1$ –, otra con un índice de autocorrelación similar a la del GS – $I \approx 0,54$ –, y una tercera con un bajo índice de Moran – $I \approx 0$ –. En la tabla 4.2. se muestra el detalle de las comunas escogidas: Quinta Normal, San Miguel y Las Condes:

<i>Área Urbana</i>	<i>Moran I</i>	<i>varianza</i>	<i>Z-score</i>	<i>p-val</i>
Quinta Normal	0,1074	$5,30 \times 10^{-5}$	14,998	<0,001
San Miguel	0,5714	$8,20 \times 10^{-5}$	63,431	<0,001
Las Condes	0,9034	$3,60 \times 10^{-5}$	150,964	<0,001
Gran Santiago	0,5475	$1,00 \times 10^{-6}$	666,888	<0,001

Tabla 4.2. Comunidades escogidas para ejecutar algoritmo de regionalización jerárquica.
Fuente: elaboración propia

Como se aprecia en la tabla, todos los valores son significativos con valor $p < 0,1\%$. Cuando se tiene un valor p estadísticamente significativo, el índice de global de autocorrelación de Moran evalúa si los patrones de la variable de estratificación escogida se pueden expresar en patrones espaciales clusterizados, cuando toma valores positivos, y más dispersos de lo que se esperaría de un patrón aleatorio si toma valores negativos. Desde esa perspectiva, las tres tipologías de comunas escogidas representan tres situaciones distintas en intensidad de autocorrelación del ISMT en el espacio, lo que permite controlar el sesgo que pudiese ocurrir debido a la distribución espacial de nuestra variable de estratificación.

En relación al nivel de clusterización en que se presentan los resultados, el criterio utilizado toma el valor de la proporción de la distancia entre los *cluster* sobre la suma total – interna y externa – de las distancias, medidas con la suma euclidiana de los cuadrados. Se escoge el nivel de clusterización que supera un valor de 0,75. Dicho valor de bondad de ajuste nos dará una primera idea de la escala en la que se agrupan los datos de SR local.

$$SS \text{ Ratio} = \frac{\text{Total between } SS}{\text{Total } SS} \geq 0,75$$

Los resultados de esta rutina para cada uno de los índices de SR estudiados se muestran en las figuras 4.4., 4.5. y 4.6., para las comunas de Quinta Normal, San Miguel y Las Condes, respectivamente.

A nivel general, se confirma que existe concordancia en relación entre el nivel de clusterización, entendido como escala de análisis, y los patrones observados en las figuras 4.1., 4.2. y 4.3., tanto entre cada índice para una misma comuna como en los resultados intercomunales para cada índice. Sin embargo, al comparar el orden de magnitud de dichas diferencias entre cada índice a nivel intra comunal, no es posible hablar de una diferencia significativa en las escalas obtenidas; sólo en el caso de Quinta Normal, la diferencia obtenida entre el índice de disimilaridad –10 clusters, agrupando entre 21 y 111 manzanas– y el índice de Theil –4 clusters, agrupando entre 100 y 252 manzanas– da cuenta de escalas de análisis distintas.

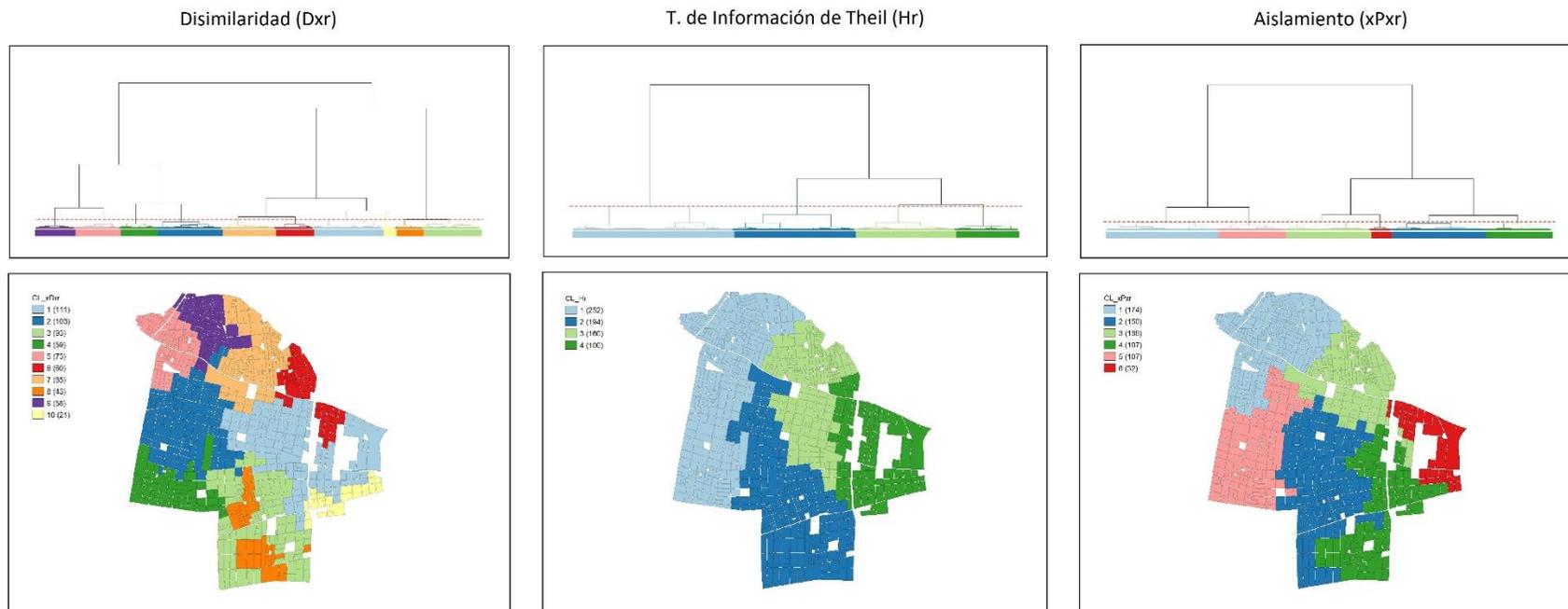


Figura 4.4.

Mapas de regionalización jerárquica y dendogramas. Quinta Normal, por manzana, 2012

Fuente: elaboración propia a partir del ISMT (OCUC, 2018)

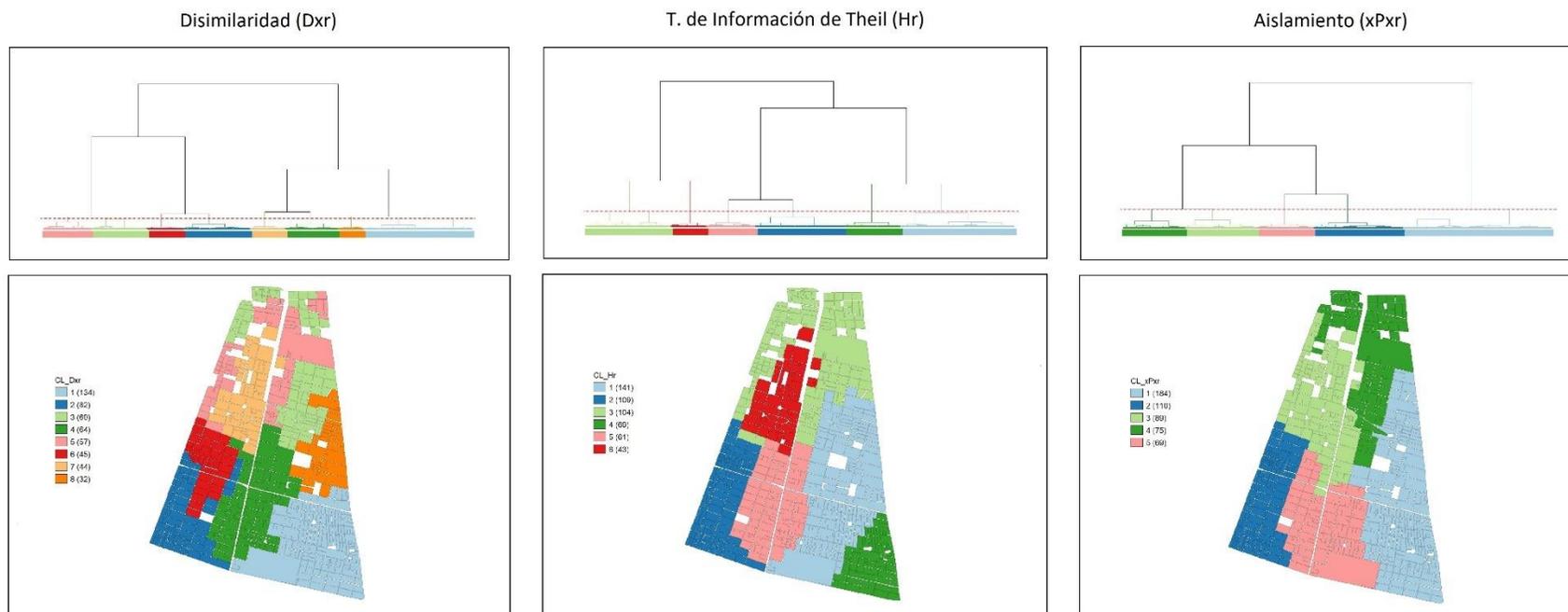


Figura 4.5.
 Mapas de regionalización jerárquica y dendogramas. San Miguel, por manzana, 2012
 Fuente: elaboración propia a partir del ISMT (OCUC, 2018)

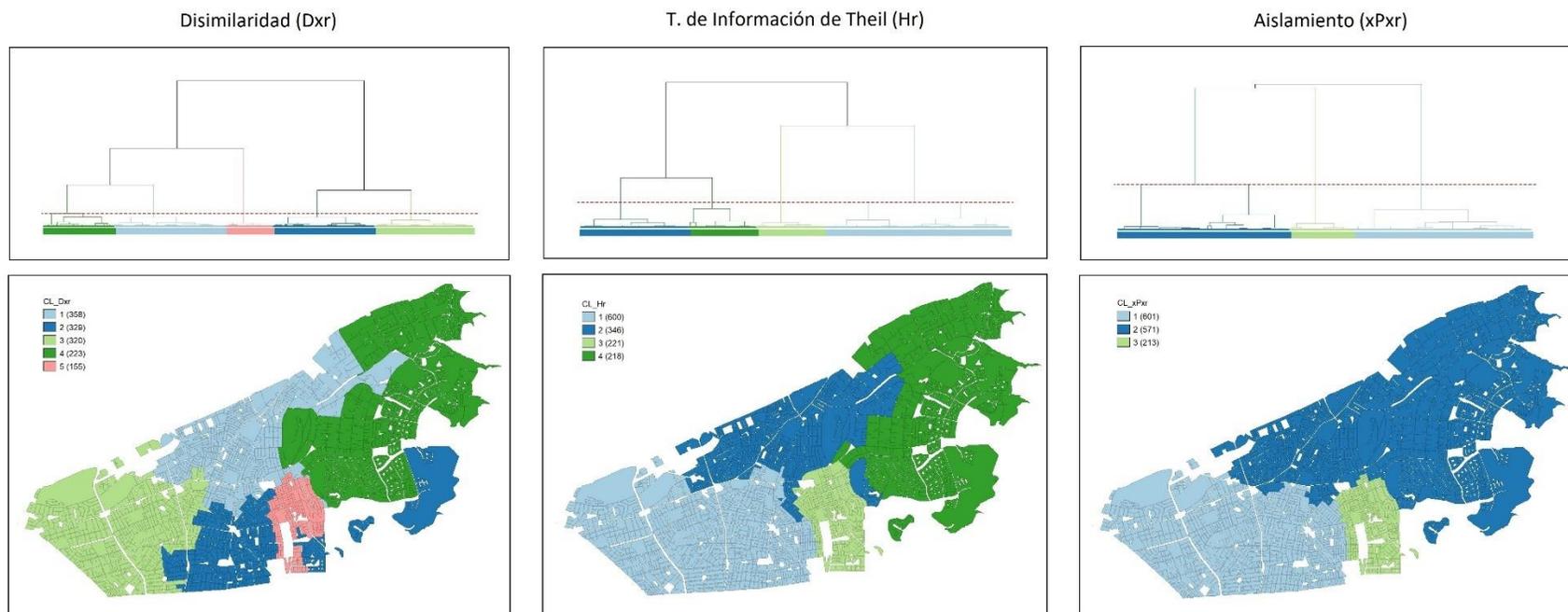


Figura 4.6.
 Mapas de regionalización jerárquica y dendogramas. Las Condes, por manzana, 2012
 Fuente: elaboración propia a partir del ISMT (OCUC, 2018)

Validación y escala óptima de análisis

Uno de los puntos más difíciles, y con menos consenso, trata sobre la validación de los *clusters* obtenidos, a saber si estos representan realmente datos agrupables por la variable medida, o bien, incluye el sesgo de agrupar “ruido” dentro de los grupos identificados (James *et al*, 2013; Garretón y Sánchez, 2016). Medir por valores de corte del *SS ratio* no otorga validez, en cuanto la disminución de la escala –aumento en el número de *clusters*– siempre trae consigo un aumento de dicho valor; en el caso extremo, en que cada observación se considera un “*cluster* de un individuo”, ese valor sería 1.

Aunque es un método subjetivo (Garretón y Sánchez, 2016), muchas veces se grafican estos valores para identificar un punto de inflexión que señale el nivel donde su aumento no sea significativo. Se muestra el gráfico correspondiente a los datos modelados para Quinta Normal y Las Condes, a fin de revisar empíricamente la dificultad que reviste identificar la escala óptima de análisis.

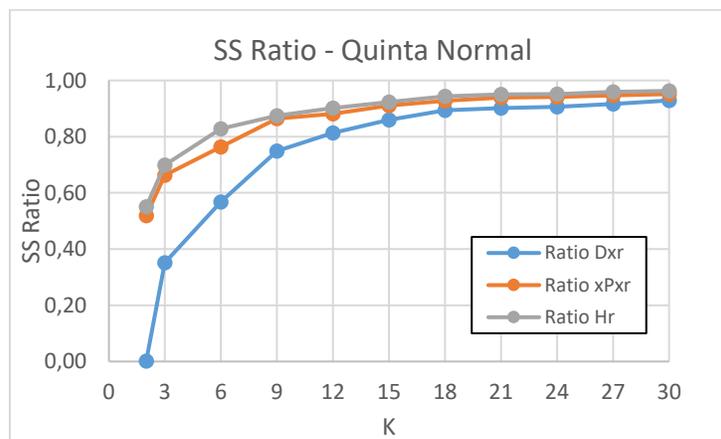


Figura 4.7.

Relación entre SS ratio y cantidad de clusters. Quinta Normal, por manzanas, 2012

Fuente: elaboración propia

En el caso de Quinta Normal, la curva de los datos de clusterización del índice de disimilaridad no presenta un claro punto de inflexión. Con un rango entre 9 y 18 clusters se hace impensable inferir una escala óptima de análisis. Del mismo modo, para los índices de aislamiento y Theil tampoco es fácil elegir un punto de corte ideal.

En Las Condes, sin embargo, aparece un caso más sencillo; la figura 4.8 muestra un punto de inflexión fácil de identificar para los datos de la modelación del índice de aislamiento. Sin embargo, esto no se refleja en los otros dos índices, donde unos saltos cualitativos en el gráfico, entre los 6 y 12 clústers, hace dudar antes de elegir el primer punto de inflexión como nivel óptimo.

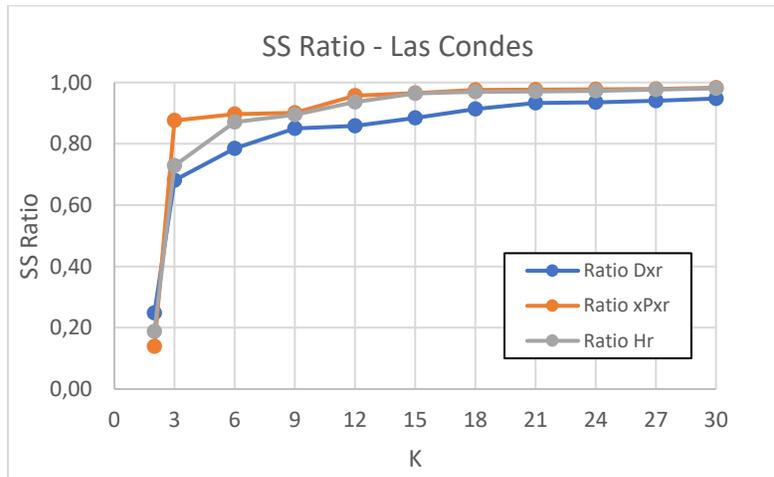


Figura 4.8.

Relación entre SS ratio y cantidad de clusters. Las Condes, por manzanas, 2012

Fuente: elaboración propia

En términos de los cambios del *SS ratio*, una correcta aproximación hacia las escalas óptimas de análisis requeriría determinar mediante pruebas estadísticas la significancia de dichos cambios. Una forma de lograrlo sería mediante el estadístico *gap*, desarrollado por Tibshirani, Walther y Hastie (2001), el cual compara la heterogeneidad interna de las particiones contra su valor esperado en una distribución nula; Garretón y Sánchez (2016) proponen comparar el valor empírico de los datos que quieren validarse con la diferencia obtenida de simulaciones de Monte Carlo, utilizando *datasets* donde las distribuciones empíricas de la base de datos original se modelan de forma aleatoria en el espacio, generando “patrones espaciales aleatorios independientes mientras se preserva la distribución estadística de cada atributo” (p.18). Estas y otras soluciones deben analizarse, a modo de encontrar la opción que mejor se adapte a la necesidad relativa a las escalas de análisis de SR.

3. Conclusiones y nuevas directrices de investigación

A partir de los resultados presentados y su respectivo análisis, sostengo lo señalado en la hipótesis de esta investigación: se observa en el GS de 2012 múltiples patrones de SR, expresados tanto en su escala local-inter barrrial, como en la meso escala de dinámicas comunales y en la macro escala metropolitana.

Bajo el lente de las dimensiones de uniformidad y exposición de los grupos socio-materiales del ISMT, es posible explorar la manera en que estos patrones de SR se entrelazan; influyendo entre sí, formando situaciones dignas de análisis en torno a las definiciones operativas de segregación.

En cuanto a los mismos índices levantados, puedo afirmar que al medirlos tomando en cuenta la espacialidad de la SR, se revaloriza su idoneidad en cuanto a la información que cada uno es capaz de mostrar en torno a las dimensiones de SR ya mencionadas.

Finalmente, las estimaciones sobre la escala óptima de análisis de SR –principal objeto de investigación de esta tesis– merecen ser relevadas alrededor de dos puntos clave.

El primero, respecto a la comparación entre índice o dimensión de segregación y su escala geográfica de análisis, medida según nivel de clusterización. En las tres comunas, independiente del nivel de autocorrelación espacial de los valores por manzana del ISMT, los *cluster* resultantes para el índice de disimilaridad son más pequeños en comparación a los índices de aislamiento y de Theil. Por su parte, los resultados casi idénticos entre estos dos últimos se condice con su similitud de patrones de composición socio-espacial, atribuible a la influencia de la dimensión de exposición por sobre la de uniformidad. Vale decir, aparecen indicios sobre la diferencia existente en su escala óptima de análisis entre ambas dimensiones de SR.

El otro aspecto a relevar guarda relación con el efecto mismo de los niveles de autocorrelación del ISMT en la región estudiada. Ante un bajo nivel de autocorrelación, la magnitud de la diferencia de escala entre dimensiones de

segregación se hace mucho más evidente; no obstante, hace mucho más complejo validar un nivel óptimo de análisis por métodos de inspección simple de los resultados del algoritmo de regionalización jerárquica. Para el caso opuesto, un alto nivel de autocorrelación de la variable de estratificación provoca una convergencia en la escala óptima de análisis de ambas dimensiones de segregación, a la vez que simplifica el análisis de validación de dicha escala.

De los resultados que tenían por objeto estimar la interacción y cobertura de las escalas geográficas específicas que toma la SR en el espacio, se desprende por un lado que estos no son concluyentes; hace falta validar de forma objetiva y significativa las escalas geográficas obtenidas. El mismo criterio corre para el efecto de la autocorrelación de la variable de estratificación, en cuanto debe estudiarse de forma particular. Si corresponde a un sesgo de resultados –problema operativo– o bien se corresponde con procesos de SR particulares, donde ambas dimensiones coinciden –fenómeno socio-espacial–, es una de las interrogantes que nacen de esta investigación.

Luego, pareciera ser que en el desarrollo de modelos de regionalización existe un lugar para la investigación en SR. Propongo, en esa línea, estudiar más en profundidad aquellos algoritmos de aprendizaje no supervisado para adaptar métodos de regionalización al estudio de SR.

A la hora de determinar y validar una escala óptima de análisis, las aproximaciones realizadas por Garretón y Sánchez (2016) son reveladoras. Sin embargo, intuyo que debe complementarse con otros métodos estadísticos de validación.

También será necesario incorporar la noción de recálculo que Garretón y Sánchez utilizan para su indicador de desintegración social a partir de análisis de componentes principales, y llevarlo hacia la construcción de índices sintéticos de SR, que logren reducir el sesgo de medición para su posterior uso comparativo, independiente del contexto temporal y espacial en las regiones metropolitanas a utilizarse.

Alrededor de las pretensiones originales de esta tesis, llegué a la idea de que, teóricamente, es posible modificar la función objetivo de los métodos de regionalización, reemplazándola por expresiones que no midan la distancia euclidiana de las variables a clusterizar, sino que midan alguna de las dimensiones de SR similar a como están operacionalizados los índices de SR clásicos. Me da la sensación de que no sólo se lograría reducir la necesidad de capacidad de cómputo de las herramientas para medir SR, sino además podría controlarse de manera consciente los sesgos por cambio de escala y nivel de agregación de los datos.

4. Reflexión final: segregación residencial y su escala

A pesar de no ser concluyentes, los indicios encontrados de que, en efecto, existe correspondencia entre distintas escalas –niveles de observación– con las dimensiones de uniformidad y exposición de SR, confirman la premisa inicial de esta investigación: para capturar la espacialidad de los fenómenos de SR, la escala geográfica debe considerarse como categoría epistemológica dentro de su estudio y medición. La implicancia práctica directa de este supuesto ha sido relevada a lo largo de esta investigación: encontrar una escala óptima de análisis para medir índices particulares –o dimensiones– de segregación. Por su parte, vale la pena revisar sus consecuencias teóricas.

Desde un punto de vista ontológico, Soja (2010, 69:71) plantea la necesidad de reestructurar la producción del conocimiento alrededor de tres cualidades o dimensiones principales de la existencia humana: social, temporal y espacial. En esa lógica, advierte Soja que “la mayor parte de nuestras teorías sociales y sus epistemologías asociadas se basan en supuestos enfocados en los aspectos sociales y temporales del ser”, en desmedro de su “espacialidad fundamental” (p.70).

En el caso particular de los estudios sobre SR, este desbalance ontológico puede ocurrir en ambas direcciones. Existen investigaciones que relevan en mayor medida los procesos de segregación que ejercen grupos sociales sobre otros en la lucha

por el espacio urbano –dimensión social y temporal–; otras investigaciones se enfocan en los patrones socio-espaciales de la segregación, aludiendo más a un “estado segregado” de la población bajo estudio, mientras se soslaya la historicidad –temporalidad– de dichos procesos. Sin embargo, el malentendido sobre las escalas de la SR proviene quizás de saber entenderla como un fenómeno multidimensional.

A partir de sub-dimensiones sociales de estratificación –distancia social– e integración/exclusión, como de sub-dimensiones espaciales de distribución geográfica –uniformidad y exposición–, se considera la SR como fenómeno socio-espacial multidimensional y multiescalar. Estas consideraciones, junto a los supuestos ontológicos recién expuestos, son afines a la definición de SR que hace Marcuse (2005), referenciada en la introducción de esta tesis. ¿Cuál es el problema entonces?

A mi modo de ver –y con relación a los resultados obtenidos de la investigación–, la principal confusión que debe cuidarse es el equiparar los índices para medir SR con el fenómeno en sí mismo. Sin miedo de ser insistente, recalco que dichos índices miden dimensiones particulares de SR según como son definidos; al tener un objeto de análisis particular, se condice la espacialidad y su escala geográfica con aquellas dimensiones involucradas.

REFERENCIAS BIBLIOGRÁFICAS

- Agostini, C.A., Hojman, D., Román, A. & Valenzuela, L. (2016). *Segregación residencial de ingresos en el Gran Santiago, 1992 - 2002: una estimación robusta*. EURE, 42 (127), 159-184.
- Anselin, L. (2018). *GeoDa Workbook*. Recuperado del sitio web de GeoDa, <https://geodacenter.github.io/documentation.html>
- Anselin, L., Syabri, I. & Kho, Y. (2006). GeoDa: An Introduction to Spatial Data Analysis [Computer software; version 1.12.1.161]. *Geographical Analysis* 38 (1), 5-22.
- Atisba Monitor (2017). Barrios críticos por narcotráfico. Recuperado del sitio web de Atisba, http://www.atisba.cl/wp-content/uploads/2017/10/Reporte-Atisba-Monitor-Barrios-Criticos_Informe.pdf
- Bailey, C.T. & Gatrell, A.C. (1995). *Interactive spatial data analysis*. Essex, UK: Longman.
- Bell, W. (1954). *A probability Model for the Measurement of Ecological Segregation*. *Social Forces*, 32 (4), 357-364.
- Borsdorf, A. (2003). *Cómo modelar el desarrollo y la dinámica de la ciudad latinoamericana*. EURE, 29 (86), 37-49.
- Bourdieu, P. (1999). Efectos de lugar. En *La miseria del mundo*, 119-124. Buenos Aires: Fondo de Cultura Económica.
- Brenner, N. (2000). *The Urban Question as a Scale Question: Reflections on Henri Lefebvre, Urban Theory and the Politics of Scale*. *International Journal of Urban and Regional Research*, 24 (2), 361-378.
- _____. (2001). *The Limits to Scale? Methodological Reflections on Scalar Structuration*. *Progress in Human Geography*, 25 (4), 591-614.
- Caldeira, T. (2007). *Ciudad de Muros*. Barcelona: Editorial Gedisa. (Trabajo original publicado en 2000).
- Castells, M. (1977). Is there an urban sociology? En Pickvance, C. (Ed.), *Urban sociology: Critical essays*, 33-59. Nueva York, NY: St. Martin's Press.
- _____. (1999). *La cuestión urbana*. (15a ed.). México: Siglo XXI. (Trabajo original publicado en 1972).
- Dotson, F. & Dotson, L. (1957). *La estructura ecológica de las ciudades mexicanas*. *Revista Mexicana de Sociología*, 19 (1), 39-66.

- Ducci, M. E. (2000). *Santiago: territorios, anhelos y temores. Efectos sociales y espaciales de la expansión urbana*. EURE, 26 (79), 5-24.
- Duncan, O.D. & Duncan, B. (1955). *A methodological analysis of segregation indexes*. American Sociological Review, 20 (2), 210-217.
- Elbers, C., Lanjouw, J. O. & Lanjouw, P. (2003). *Micro-level estimation of poverty and inequality*. Econometrica, 71 (1), 355-364.
- ESRI (2017). ArcGIS Desktop: Release 10.5.1 [Computer Software]. Redlands, CA, EE.UU.: Environmental Systems Research Institute.
- Fowler, C.S. (2016). *Segregation as a multiscalar phenomenon and its implications for neighborhood-scale research: the case of South Seattle 1990–2010*. Urban Geography, 37 (1), 1-25.
- Galster, G.& Killen, S. (1995) *The Geography of Metropolitan Opportunity: A Reconnaissance and Conceptual Framework*. Housing Policy Debate, Metropolitan Institute at Virginia Tech, 6 (1), 7-43.
- Garretón, M. & Sánchez, R. (2016). *Identifying an optimal análisis level in multiscalar regionalization: A study case of social distress in Greater Santiago*. Computers, Environment and Urban Systems, 56, 14-24.
- Garrocho, C. & Campos-Alanís, J. (2013). *Réquiem por los indicadores no espaciales de segregación residencial [SR]*. Papeles de Población, Julio-Septiembre, 269-300.
- Hall, P. (1996). *Ciudades del mañana: historia del urbanismo en el siglo XX*. Barcelona: Serbal.
- Halbwachs, M. (2004). *Chicago, experiencia étnica*. Reis, 108, 215-253. (Trabajo original publicado en 1932).
- Harvey, D. (1989). *The urban experience*. Baltimore, MD: The John Hopkins University Press.
- Hernández, R., Fernández, C. y Baptista, M. (2010). *Metodología de la investigación*. (5ª ed.). México: McGraw Hill.
- Hester, J. (2018). glue: Interpreted String Literals. R package version 1.3.0. Recuperado de <https://CRAN.R-project.org/package=glue>
- Hidalgo, R. (2004). *De los pequeños condominios a la ciudad vallada: Las urbanizaciones cerradas y la nueva geografía social en Santiago de Chile (1990-2000)*. EURE, 30 (91), 29-52.
- James, G., Witten, D., Hastie, T. & Tibshirani, R. (2013). *An Introduction to Statistical Learning, with Applications in R*. New York, NY: Springer-Verlag.

- Jencks, C. & Mayer, S. (1990). The Social Consequences of Growing Up in a Poor Neighborhood. En L.E. Lynn & M. McGeary (Eds.), *Inner-City Poverty in the United States* (pp. 111-186). Washington, D.C.: National Academy Press.
- Jones, K. (1998). *Scale as epistemology*. *Political Geography*, 17 (1), 25-28.
- Lee, B.A., Firebaugh, G., Matthews, S.A., Reardon, S.F., Farrell, C.R., & O'Sullivan, D. (2008). *Beyond the census tract: Patterns and determinants of racial residential segregation at multiple scales*. *American Sociological Review*, 73 (5), 766–791.
- Lee, C. & Culhane, D. (1998). *A perimeter-based clustering index for measuring spatial segregation: a cognitive GIS approach*. *Environment and Planning B: Planning and Design*, 25 (3).
- Lefebvre, H. (1969). *El derecho a la ciudad*. Barcelona: Península.
- Lefebvre, H. (2013). *La producción del espacio*. Madrid: Capitán Swing Libros. (Trabajo original publicado en 1974).
- Lencioni, S. (2015). Totalidad y tríadas: comprendiendo el pensamiento de Lefebvre. En C.A. de Mattos y F. Link (Eds.), *Lefebvre revisitado: capitalismo, vida cotidiana y el derecho a la ciudad*. Santiago: RIL editores – Instituto de Estudios Urbanos y Territoriales UC.
- Lieberson, S. & Carter, D.K. (1982). *A Model for Inferring the Voluntary and Involuntary Causes of Residential Segregation*. *Demography* 19 (4), 511-526.
- Lojkin, J. (1986). *El marxismo, el estado y la cuestión urbana*. México: Siglo XXI. (Trabajo original publicado en 1977).
- Marcuse, P. (2005). Enclaves yes, ghettos no. En D. Varady (Ed.), *Desegregating the city: Ghettos, enclaves, and inequality*, 15-30. Albany, NY: State University of New York Press.
- Marston, S. (2000). *The social construction of scale*. *Progress in Human Geography*, 24 (2), 219-242.
- Marston, S., Jones, J.P. & Woodward, K. (2005). *Human geography without scale*. *Transactions of the Institute of British Geographers*, 30 (4), 416-432.
- Marston, S. & Smith, N. (2001). *States, scales and households: Limits to scale thinking? A response to Brenner*. *Progress in Human Geography*, 25 (4), 615-619.
- Massey, D. B. (1994). *Space, place, and gender*. Minneapolis, MN: University of Minnesota Press.
- Massey, D. B. (2005). *For space*. London: SAGE.

- Massey, D. S. & Denton, N. (1988). *The dimensions of residential segregation*. Social Forces, vol. 67, n° 2.
- Monkkonen, P. & Zhang, X. (2014). *Innovative measurement of spatial segregation: Comparative evidence from Hong Kong and San Francisco*. Regional Science and Urban Economics, 47, 99-111.
- Observatorio de Ciudades UC (2018). Metodología para la construcción del Índice Socio Material Territorial [Documento de Trabajo]. Recuperado de www.observatoriodeciudades.com
- Openshaw, S. (1984). *The modifiable spatial unit problem*. Concepts and Techniques in Modern Geography, n° 38. Norwich, England: Geo Books.
- Östh, J., Clark, W.A.V. & Malmberg, S. (2015). *Measuring the scale of segregation using k-nearest neighbor aggregates*. Geographical Analysis, 47, 34–49.
- Östh, J., Malmberg, B. & Andersson, E.K. (2014). Analysing segregation using individualized neighbourhoods. En C.D. Lloyd, I.G. Shuttleworth & D.W.S. Wong (Eds.), *Socio-spatial segregation: Concepts, processes and outcomes*, 135–162. Bristol, UK: The Policy Press.
- Park, R. & Burgess, E. (1967). *The City*. Chicago, IL: The University of Chicago Press. (Trabajo original publicado en 1925).
- Purcell, M. (2003) *Islands of practice and the Marston/Brenner debate: Toward a more synthetic critical human geography*. Progress in Human Geography, 27 (3), 317-332.
- R Core Team (2017). foreign: Read Data Stored by 'Minitab', 'S', 'SAS', 'SPSS', 'Stata', 'Systat', 'Weka', 'dBase', R package version 0.8-70. Recuperado de <https://CRAN.R-project.org/package=foreign>
- R Core Team (2018). R: A language and environment for statistical computing [Computer software]. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Rasse, A. (2015). *Juntos pero no revueltos. Procesos de integración social en fronteras residenciales entre hogares de distinto nivel socioeconómico*. EURE, 41 (122), 125-143.
- Reardon, S.F. & Firebaugh, G. (2002). *Measures of multigroup segregation*. Sociological Methodology, 32, 33-67.
- Reardon, S.F. & O'Sullivan, D. (2004). *Measures of spatial segregation*. Sociological Methodology, 34, 121-162.

- Rodríguez, G. (2013) *El uso de zonas censales para medir la segregación residencial. Contraindicaciones, propuesta metodológica y un estudio de caso: Argentina 1991-2001*. EURE, vol. 39, n° 118, pp. 97 – 122.
- Rodríguez, J. (2001) *Segregación residencial socioeconómica: ¿qué es?, ¿cómo se mide?, ¿qué está pasando?, ¿importa?* Santiago de Chile: Serie Población y desarrollo N° 16, CEPAL.
- Ruiz-Tagle, J. (2016) *La persistencia de la segregación y la desigualdad en barrios socialmente diversos: un estudio de caso en La Florida, Santiago*. EURE, 42 (125), 81-108.
- Ruiz-Tagle, J. & López, E. (2014) *El estudio de la segregación residencial en Santiago de Chile: revisión crítica de algunos problemas metodológicos y conceptuales*. EURE, 40 (119), 25-48.
- Sabatini, F., Cáceres, G. & Cerda, J. (2001). *Segregación residencial en las principales ciudades chilenas: Tendencias de las tres últimas décadas y posibles cursos de acción*. Revista EURE, vol. 27, n° 82.
- Sabatini, F., Cáceres, G. & Rasse, A. (2013). Bifurcación de senderos: Entre la segregación que “guetiza” los barrios populares y la gentrificación que ayuda a su “moyenización”. En F. Sabatini, G. Wormald & A. Rasse (Eds.), *Segregación de la vivienda social: ocho conjuntos en Santiago, Concepción y Talca*, 221-242. Santiago: Colección Estudios Urbanos UC.
- Sabatini, F. Sierralta, (2006). *Medición de la segregación residencial: meandros teóricos y metodológicos y especificidad latinoamericana*. En J.M. Pinto da Cunha (Ed.), *Novas Metrópoles Paulistas; População, Vulnerabilidade e Segregação*. Campinas: Nepo-Unicamp.
- Sabatini, F.; Wormald, G.; Sierralta, C. & Peters, P. (2010). *Segregación residencial en Santiago: Tendencias 1992-2002 y efectos vinculados con su escala geográfica*. En F. Sabatini, R. Salcedo, G. Wormald & G. Cáceres (Eds.), *Tendencias de la segregación en las principales ciudades chilenas: Análisis censal 1982-2002*, 19-42. Santiago: Pontificia Universidad Católica de Chile/Instituto Nacional de Estadísticas.
- Sampson, R. (2012). *Great American city: Chicago and the enduring neighborhood effect*. Chicago, IL: University of Chicago Press.
- Soja, E. (2008). *Postmetrópolis. Estudios críticos sobre las ciudades y las regiones*. Madrid: Traficantes de sueños.
- Soja, E. (2010). *Seeking Spatial Justice*. Minneapolis, MN: University of Minnesota Press.

- Theil, H. (1972). *Statistical decomposition analysis, with applications in the social and administrative sciences*. Amsterdam: North-Holland.
- Tibshirani, R., Walther, G. & Hastie, T. (2001). *Estimating the number of clusters in a data set via the gap statistic*. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 63(2), 411–423.
- Tilly, C. (2000). *La desigualdad persistente*. Buenos Aires: Manantial
- Wacquant, L. (2007). *Los Condenados de la Ciudad. Gueto, Periferias y Estado*. Buenos Aires.: Siglo XXI Editores.
- Wacquant, L. (2009). *The body, the ghetto and the penal state*. Qualitative Sociology, 32 (1),101-129.
- Wagensberg, J. (2004). *La rebelión de las formas. O cómo perseverar cuando la incertidumbre aprieta*. Barcelona: Tusquets.
- White, M. (1983). *The measurement of spatial segregation*. American Journal of Sociology, 88 (5).
- Wickham, H. (2017). tidyverse: Easily Install and Load the 'Tidyverse'. R package version 1.2.1. Recuperado de <https://CRAN.R-project.org/package=tidyverse>
- Wilson, W.J. (1987). *The truly disadvantaged: The inner city, the underclass, and public policy*. Chicago, IL: University of Chicago Press.
- Wirth, L. (1962). *El urbanismo como modo de vida*. Buenos Aires: Ediciones 3. (Trabajo original publicado en 1938).
- Wong, D.W. (2003). *Spatial decomposition of segregation indices: A framework toward measuring segregation at multiple levels*. Geographical Analysis, 35 (3), 179–194.
- Wong, D.W. (2004). *Comparing traditional and spatial segregation measures: a spatial scale perspective*. Urban Geography, 25 (1).