

DETERMINANTES SOCIOESPACIALES DEL EMPLEO EN CALI

SEBASTIÁN CUADROS CARDOZO

**UNIVERSIDAD DEL VALLE
FACULTAD DE CIENCIAS SOCIALES Y ECONOMICAS
DEPARTAMENTO DE ECONOMIA
SANTIAGO DE CALI
2012**

DETERMINANTES SOCIOESPACIALES DEL EMPLEO EN CALI

**SEBASTIÁN CUADROS CARDOZO
CODIGO 0638948**

**DIRIGIDO POR:
HARVY VIVAS PACHECO**

**UNIVERSIDAD DEL VALLE
FACULTAD DE CIENCIAS SOCIALES Y ECONOMICAS
DEPARTAMENTO DE ECONOMIA
SAN TIAGO DE CALI
2012**

CONTENIDO

	pág.
INTRODUCCIÓN	1
1. REVISIÓN DE LITERATURA RELACIONADA	2
2. MARCO TEÓRICO	6
2.1 LOS AVANCES GENERALES DE LA GEOGRAFÍA DE LAS OPORTUNIDADES EN LOS BARRIOS SEGREGADOS.....	7
2.2 SEGREGACIÓN RESIDENCIAL Y MODELO DE DISTANCIA SOCIAL	8
3. HETEROGENEIDAD EN LAS PREFERENCIAS	12
3.1 COSTOS DE INTERACCIÓN	12
3.2 ENTORNO SOCIAL: RAZA EDUCACIÓN Y OCUPACIÓN	12
3.3 CLUSTERS ESPACIALES	13
4. DATOS.....	16
5. ANÁLISIS EXPLORATORIO DE DATOS ESPACIALES	17
5.1 ANÁLISIS UNIVARIADO	17
5.2 ANÁLISIS BIVARIADO.....	23
6. ANÁLISIS CONFIRMATORIO DE DATOS ESPACIALES	27
6.1 MODELO DE REGRESIÓN ESPACIAL CON EXTERNALIDADES GLOBALES	27
6.2 ESTIMACIONES.....	30
7. CONCLUSIONES.....	36
REFERENCIAS BIBLIOGRÁFICAS	38
ANEXOS	41

LISTA DE TABLAS

	pág.
Tabla 1. Medida de autocorrelación global I de Morán de la variable de capital humano estimado	27
Tabla 2. Test LM y LM robusto de EDUESTIMAD	28
Tabla 3. Estimaciones primera etapa por MCO	30
Tabla 4. Estimaciones segunda etapa.....	32

LISTA DE FIGURAS

	pág.
Figura 1. Diagrama de caja y su representación gráfica de la variable porcentaje de afrocolombianos por barrio	18
Figura 2. Diagrama de caja y su representación gráfica de la variable promedio de edad por barrio	19
Figura 3. Diagrama de caja y su representación gráfica de la variable porcentaje de personas que tienen nivel académico superior y postgrado.	19
Figura 4. Índice Global de Morán y mapa de significancia de clusters para la variable porcentaje de afrocolombianos por barrio	21
Figura 5. Índice Global de Morán y mapa de significancia de clusters para la variable promedio de edad por barrio	¡Error! Marcador no definido.21
Figura 6. Índice Global de Morán y mapa de significancia de clusters para la variable porcentaje de personas que tienen nivel académico superior o postgrado.....	22
Figura 7. Índice Global bivariado de Morán y mapa de significancia de clusters para las variables porcentaje de población afrocolombiana y porcentaje de personas de 25 años o más que tienen nivel académico superior o postgrado.....	24
Figura 8. Índice Global bivariado de Morán y mapa de significancia de clusters para las variables porcentaje de población afrocolombiana y edad promedio por barrio.....	25
Figura 9. Índice Global bivariado de Morán y mapa de significancia de clusters para las variables porcentaje de personas de 25 años o más que tienen educación superior o posgrado y edad promedio en los barrios.....	26

Figura 10. Dispersión de Morán para los residuales del modelo de rezago espacial.....	34
Figura 11. Test de permutaciones	35

DETERMINANTES SOCIO-ESPACIALES DEL EMPLEO EN CALI¹

SEBASTIÁN CUADROS CARDOZO

RESUMEN

Este trabajo estudia los patrones de desempleo en la ciudad de Cali con información del censo de 2005 a nivel de barrios. Para ello se estudian tanto las características del entorno al interior de cada barrio capturadas por la composición racial, un índice de capital humano y la edad promedio que se aproxima a la estructura por edades, como las características de los barrios vecinos y cómo estas impactan sobre la distribución y aglomeración del desempleo aproximado por la tasa de subutilización de cada barrio. Teniendo como bases teóricas la geografía de las oportunidades de Flores (2008) y el modelo de distancia social de Akerlof (1997), se emplea como herramienta un modelo en dos etapas, la primera mediante MCO y la segunda mediante un modelo de rezago espacial. Los resultados sugieren que el entorno de cada barrio y el efecto de sus vecinos impactan significativamente sobre el desempleo en cada uno de ellos.

Palabras clave: Modelos espaciales

Clasificación JEL: C21

¹ Esta investigación se desprende de una investigación más amplia denominada “Segregación residencial con econometría espacial en Cali” bajo la dirección del Dr. Harvy Vivas Pacheco y el apoyo de la Vicerrectoría de investigaciones, Universidad del Valle.

INTRODUCCIÓN

La configuración de las características de la población en la ciudad de Cali evidenciada en la segregación residencial está estrechamente relacionada con el factor migración. Uno de los principales flujos migratorios proveniente de la costa pacífica se estableció en su mayoría en barrios del oriente de la ciudad, mientras que en la zona de ladera del occidente se localizaron otros importantes grupos migratorios originarios del altiplano andino de los departamentos del Valle, Cauca y Nariño.

Esta polarización demográfica del espacio urbano se asocia a las restricciones y preferencias de localización de los individuos y su tendencia a establecerse cerca de personas con características similares como en este caso, el origen, que facilita el intercambio social. Otro elemento que reproduce esta polarización es la pobre condición social y económica con la que llega la población inmigrante que limitan sus posibilidades de localización hacia zonas donde el precio del suelo es bajo, que corresponden a las zonas oriental y de ladera occidental.

El principal objetivo de este estudio es confirmar mediante técnicas propias de la econometría espacial, la hipótesis de que la segregación de la población ha generado distintos entornos socioeconómicos al interior de los barrios de la ciudad y que estos se ven reflejados en las diferencias observadas del desempleo y la subutilización de la mano de obra. El estudio enfatiza en las concentraciones de comunidades afrocolombianas que representan más del 26% de la población total, siendo quizá el más alto de las principales ciudades del país.

Al respecto son pocos los estudios en la región que han tratado de identificar la influencia del entorno socioeconómico sobre la situación laboral, por lo que este trabajo representa un valioso aporte a este campo de estudio.

1. REVISIÓN DE LITERATURA RELACIONADA

La investigación de la influencia espacial sobre las relaciones sociales interpersonales, encuentra que la distancia de los participantes es una variable significativa. Del mismo modo que las empresas tienen geográficamente incentivos de permanecer cerca de sus clientes actuales, los individuos también tienen una buena razón para no abandonar sus parientes y amigos.

Akerlof (1997) plantea un modelo de *distancia social* en el que estudia el proceso de intercambio social. Este modelo permite identificar la presencia de subgrupos con normas y valores característicos que se comportan de manera diferente de la mayoría de la población. Sostiene además que el rasgo más importante de estos subgrupos en relación con la distribución en el espacio es un hecho puramente comercial que afecta el beneficio de los individuos pertenecientes a cada grupo y que se incrementará con la proximidad en el espacio.

Duncan y Duncan (1955) ya habían observado esta tendencia de los individuos a localizarse cerca de otros con características similares, estableciendo así una estrecha relación entre distancia social y espacial, próximo al concepto de segregación residencial utilizado con frecuencia en el campo específico de la economía regional y urbana. Empíricamente observan que las dinámicas de segregación en las ciudades, conducen a la existencia de una relación directa entre centralización residencial y situación socioeconómica, descubriendo análogamente que en la periferia urbana se localizan los individuos de menor estatus socioeconómico.

Wheeler y Stutz (1971) quisieron evaluar las hipótesis que implícitamente incorpora la relación entre distancia social y espacial, la primera es que la interacción de los individuos aumenta a medida que la distancia que los separe sea menor. Y la otra, es que las personas tienden a elegir como contactos sociales, a individuos con un estatus socioeconómico similar al suyo. Para ello emplearon datos para el año 1965 del área metropolitana de Lansing (capital del estado de Michigan) haciendo uso de las distancias métricas entre cada una de las calles de la ciudad y la información de una muestra de individuos con vehículos, acerca de los viajes puramente sociales (visitas familiares o a amigos). Los autores encontraron que en su mayoría, los viajes en vehículo realizados por los individuos, estaban inversamente relacionados con la distancia entre vías que lo separan de su destino, dedujeron también a partir de este resultado, que la distancia tiene una mayor incidencia negativa en los viajes a pie.

También examinaron las composiciones sociales dentro de los barrios y observaron que las personas

con determinados tipos de trabajo, presentaban mayor interacción social con individuos que desempeñan las mismas actividades; confirmando ambos resultados las hipótesis iniciales. A propósito del tema laboral, en un trabajo empírico Topa (2001) estima un modelo de interacciones locales en el mercado laboral, asumiendo que los agentes intercambian información con sus contactos sociales acerca de oportunidades de trabajo. En síntesis, el documento se plantea para demostrar que la distancia física es un determinante importante para la formación de redes de los individuos.

Empleando una metodología con áreas censales como lo son las vías con datos en la ciudad de Chicago, encuentra que el desempleo se caracteriza por una alta correlación espacial. Otro resultado importante, es que cuando hay alta diversidad étnica, este contagio local es menos fuerte en comparación cuando son grupos étnicamente homogéneos. Por último, se encuentra también que estos efectos secundarios son mayores con niveles de educación más bajos. En un artículo similar, Conley y Topa (2002) examinan más a fondo los patrones de desempleo en Chicago entre los años 1980 y 1990, incluyendo la relación entre su aglomeración contra distintos indicadores socioeconómicos que reflejan la distancia de las estructuras de las redes sociales de los agentes, específicamente, emplean la distancia física, el tiempo de viaje y la distribución tanto étnica y racial, como la ocupacional de cada subgrupo de la ciudad.

Su objetivo es determinar si la situación laboral de cada individuo está relacionada con la información que intercambia localmente con su red social. Encuentran que cuando las combinaciones de desplazamiento físico y tiempo de viaje, se utilizan junto con la distancia étnica y racial (distancia euclidiana en porcentajes de los tipos de etnias y razas entre los subgrupos), este último indicador parece ser la unidad de dependencia espacial que más influye en el desempleo. Otro factor que influye indirectamente sobre el desempleo es el estatus socioeconómico que tiene una gran incidencia sobre los logros educativos, no sólo por el hecho de que las personas de más bajo estatus cuentan con menos recursos escolares (libros, computadores) sino principalmente como lo sugieren Considine y Zappalà (2002), por el cúmulo de aspiraciones y expectativas que tienen estas personas derivado del entorno en que viven que crean en ellas actitudes negativas hacia el estudio por lo que invierten menos en educación lo que limita el acceso exitoso al mercado laboral. Este proceso de interacción social, como lo plantea Akerlof *op. cit.* genera externalidades que influyen no sólo en la decisión de cuanta educación demandar sino en otro tipo de decisiones sociales como: la práctica de la discriminación, la decisión de casarse, de divorciarse o de tener hijos, la decisión de si se cometen o no delitos, etc. Así, el autor expone la diferencia entre las decisiones sociales y económicas; donde la principal divergencia,

es que las primeras tienen implicaciones sociales, mientras que las segundas afectan al individuo que determina su accionar.

Sobre los factores que influyen en la demanda educativa, y siguiendo el planteamiento de Considine y Zappalà (2002), Casas et al (2002) sostiene que las personas con bajo nivel socioeconómico se enfrentan con oportunidades limitadas pues no cuentan con los recursos suficientes para acceder a una educación de calidad. Por esta misma línea Vivas (2008) encuentra que en Colombia existe segregación educativa por nivel socioeconómico, los más pobres asisten principalmente a colegios públicos y los que no son considerados pobres asisten en su mayoría a colegios privados, que tienen una mayor calidad medida como el promedio de resultados en las pruebas nacionales.

Un caso particular que se desprende del resultado de que los individuos presenten mayor interacción con individuos similares es la segregación residencial por grupos étnicos, Bayer et al (2001) argumentan que para entender las fuerzas que inciden sobre ella, es necesario otorgar a los hogares el supuesto de heterogeneidad en las preferencias de un conjunto de múltiples opciones existentes en el campo de elección, incluyendo atributos sobre la localización, escogencia de las escuelas, la delincuencia, servicios ambientales, la composición socio demográfica, características de la vivienda y el precio de la misma; y muestran cómo estas preferencias varían de acuerdo a características propias de un hogar, como la raza, la educación, los ingresos y el lugar de trabajo. La segregación residencial por grupos raciales, es explicada en gran proporción tanto por las consideraciones sobre el nivel de oferta de bienes públicos locales, como por las preferencias de los hogares sobre la raza de sus vecinos, una concepción similar a la mirada de la distancia social, donde los agentes buscan minimizarla al localizarse en este caso cerca de miembros de su misma raza. Por otro lado, si un hogar se localiza cerca de escuelas de acuerdo con su disposición a pagar por la calidad, el entorno escolar se organizaría en función de los ingresos de los hogares, generando probablemente un fenómeno de segregación residencial que podría no solo explicarse por las diferencias en los ingresos sino también como encuentran Cutler y Glaeser (1997), por las diferencias de razas, quienes en su estudio para Estados Unidos, demuestran la existencia de correlación positiva entre el porcentaje de población negra que vive en áreas segregadas y menores ingresos.

Borjas (1995) recurre a este enfoque y plantea que el efecto de la raza y la etnia sobre la acumulación de capital humano mediante el proceso de transmisión intergeneracional, resume el impacto de las características del vecindario, homogéneas para los residentes del barrio. La composición racial en

ciertos entornos, constituye un actor económico, social y cultural de mayor importancia, cuanto mayor sea el grado de segregación del barrio; después de todo, la experiencia e interacción social, están fuertemente influenciadas y determinadas con los contactos más frecuentes. Según Forero (2007), estas vecindades étnicas y raciales, provocan una relación entre posicionamiento geográfico y distintos niveles de calidad de capital humano; los individuos intercambian elementos de su contexto educativo, social y económico incorporando un conjunto de actitudes y capacidades culturales, contactos sociales y oportunidades económicas que afectan directamente su nivel como trabajadores productivos a través de su vida laboral.

2. MARCO TEÓRICO

Con frecuencia para analizar la segregación residencial se recurre a la hipótesis de *spatial mismatch* o desemparejamiento espacial planteada por Kain (1968) la cual sostiene que la concentración de la pobreza y el desempleo en su mayoría en individuos negros en determinados barrios de las ciudades es explicada en parte por el aislamiento geográfico de estos barrios a los lugares de trabajo.

Surgen tres ideas a partir de esta hipótesis donde la primera sustenta que la segregación residencial de los afros en las áreas metropolitanas no es voluntaria, sino que es en gran parte el resultado de la discriminación racial en el mercado de vivienda. Este aislamiento genera mayores desplazamientos que son costosos para los negros y reduce la información sobre oportunidades de trabajo ocasionando que estos individuos sean más propensos a trabajar cerca de sus viviendas o en ellas mismas. Otro factor identificado por Kain que puede reforzar esta tendencia es la posibilidad de que los negros experimenten menos discriminación por parte de los consumidores en áreas donde representan un mayor porcentaje del total de la población residente debido a que los clientes blancos podrán tener una aversión al tratar con los empleados negros, provocando que los empleadores contraten menos negros en áreas predominantemente blancas.

La segunda idea detrás de la hipótesis de *spatial mismatch* es causada por la discriminación que limita las zonas residenciales de los negros ocasionando que su conjunto de oportunidades de trabajo sea menor en relación a las oportunidades de los blancos, donde sus decisiones de localización dependen de las preferencias sobre los precios y los ingresos. Un limitado conjunto de oportunidades da como resultado un aumento en el desempleo.

La última idea que se desprende se refiere a que existe un desajuste espacial en el sentido de que hay un excedente de trabajadores en relación con el número de empleos disponibles en las zonas donde se concentra la población negra y una escasez de mano de obra blanca en relación a los puestos de trabajo que reside fuera de estas áreas. El excedente de mano de obra de individuos negros se traducirá en una mayor tasa de desempleo si los salarios son inflexibles. Si por el contrario los salarios son flexibles, el excedente de trabajo se traducirá en un menor nivel en los salarios de equilibrio. También es posible que algunos trabajadores negros que no pueden encontrar empleo dentro del gueto sean capaces de viajar a trabajos más distantes sufriendo de esta forma una pérdida de bienestar que se traduce en un menor salario neto. Así, la segregación de trabajo combinado con la segregación residencial puede reducir el bienestar económico de los negros en relación a las áreas donde se localizan las mayorías de

blancos.

En este estudio sin embargo no se empleará este enfoque tradicional debido a la limitación en la información desagregada de datos sobre puestos de trabajo en la ciudad de Cali y se optará por otras perspectivas teóricas como lo son el modelo de distancia social y la geografía de las oportunidades.

2.1 LOS AVANCES GENERALES DE LA GEOGRAFÍA DE LAS OPORTUNIDADES EN LOS BARRIOS SEGREGADOS

La segregación residencial se asocia a veces con el aislamiento social. Tal y como plantea Flores (2008), una situación en la que los residentes de las zonas segregadas están menos expuestos a las normas y a las oportunidades de la sociedad en general, se caracteriza principalmente por una sociedad aislada socialmente. Esto a su vez repercute en la socialización y formación de los niños. Las altas tasas de desempleo de los adultos, el rápido crecimiento de las familias monoparentales, las altas tasas de encarcelamiento, algunas conductas como el alcoholismo, la drogadicción y la violencia; tienen un efecto profundo en la normativa, en la información y en la calidad de las oportunidades locales disponibles para los niños.

La teoría se centra en el distanciamiento cultural entre las áreas concentradas de pobreza y el resto de la sociedad, que sin duda influye negativamente en los resultados educativos de los individuos, sabiendo que la educación formal se proporciona en mayor medida para las clases medias y altas. Es por ello que los individuos de estas zonas pobres y segregadas, están culturalmente en desventaja, ya que las cosas que aprenden en sus hogares y barrios, por lo general van en contra de las necesidades de la educación formal, no solo por la baja calidad de la educación, sino también porque los mecanismos de socialización de estos barrios no son los apropiados para el éxito en la escuela.

La construcción de un marco teórico que consiste en combinar los resultados educativos y el efecto del entorno socioeconómico que impacta la cantidad y calidad de oportunidades disponibles en los individuos, requiere un esfuerzo combinado de diferentes disciplinas, tales como la economía y la sociología urbana. Literatura que ayuda a comprender los efectos de la estructura y cultura en los barrios sobre el desarrollo y logros educativos de los individuos y cuál es el efecto de éstos sobre las oportunidades laborales.

Puede argumentarse que las oportunidades de educación y características sociales en los límites de un barrio, difieren de las de otros barrios, así como también debe decirse que la relación entre los entornos

de los barrios y los resultados educativos no es casual, de hecho, una alta proporción en las diferencias en la educación, están explicadas por la desigualdad entre los barrios; así como también explica por qué los resultados en las escuelas de los individuos y sus propias cualidades y la de sus familias, están tan estrechamente relacionadas en el espacio. Además hay que reconocer que no solo existe relación dentro de los barrios, sino que también con las características de los barrios circundantes. Estos dos procesos espaciales, intra-barrio e inter-barrio, ayudan a comprender la relación espacial entre pobreza y los resultados educativos, conocidos en la literatura como la heterogeneidad espacial y dependencia espacial, que también es conocida como externalidades espaciales.

2.2 SEGREGACIÓN RESIDENCIAL Y MODELO DE DISTANCIA SOCIAL

Hasta este momento se ha centrado la atención en la relación entre el entorno socioeconómico que determina en gran medida las oportunidades de los individuos a través del efecto que representa para los distintos resultados educativos de los barrios. Falta pues, hacer énfasis en otra característica propia de cada barrio y que genera grandes externalidades tanto en la formación de diferencias de capital humano, como en el estatus socioeconómico dentro de los mismos; la conformación étnica y racial de cada uno. Aspectos propios empleados en la Teoría de la Segregación Residencial y más exactamente, en la segregación residencial racial.

En los modelos de segregación residencial, se introduce el supuesto de heterogeneidad en las preferencias de los hogares para argumentar la existencia de subgrupos en la población con sus propias normas y valores, consecuencia importante de la teoría de la interacción social, que explica cómo la relación entre los individuos puede generar esos grupos que se pueden interpretar como subculturas o clases sociales estables. En el modelo de distancia social de Akerlof; *op. cit.*, se emplea este enfoque incorporando una variable de control en los hogares, la decisión de localizarse en lugares de acuerdo con sus preferencias sobre espacios sociales, que representa una aproximación al comercio local y beneficia a los individuos pertenecientes a los mismos y donde sus efectos aumentan en función de la proximidad en este espacio. Además se añade una variable de estado, el estatus social de cada individuo, ya que es un aspecto heredado; se concluye que ambas variables afectan las interacciones sociales que reflejan los aspectos locales. Para cerrar el modelo se supone que los individuos tienen expectativas estáticas sobre la localización que elegirán el resto de individuos en el espacio social.

Mediante las anteriores especificaciones, el modelo puede predecir trampas de bajo nivel con equilibrios estables en distintos subgrupos, debido a que los individuos que presentan una baja herencia

social, preferirán localizarse en zonas con características similares. Es entonces un comportamiento conformista, pues aquellos que presentan homogéneas cualidades, están más propensos a interactuar, provocando una baja posibilidad para que exista movilidad e interacción social con grupos e individuos heterogéneos. El modelo da como resultado endógeno el concepto de rendimientos de proximidad social en la dimensión de la ubicación social y geográfica.

El modelo se auto-refuerza y conduce al equilibrio estable si se le introduce beneficios negativos sobre la utilidad, pues los individuos para minimizar el riesgo de provocar envidia sobre vecinos, amigos y hasta incluso familiares, tienen incentivos a localizarse en esas zonas donde se posicionan individuos con similares características. Lo ingenioso del modelo radica en que añade elementos propios del comercio internacional y las leyes de la física sobre la gravedad para aproximar el intercambio social y lograr generar una forma funcional para los beneficios de la interacción entre los individuos. El modelo de gravedad para explicar la interacción comercial entre dos países, predice que la actividad mutua entre ellos, dependerá proporcionalmente del PIB de cada uno e inversamente proporcional al cuadrado de la distancia física de los mismos. Esta ley de la gravedad relaciona la fuerza de atracción, que es inversamente proporcional al cuadrado de la distancia entre dos cuerpos y aumenta sin límite cuando la distancia se acerca a cero. La implicación de la adopción de esta forma funcional dentro del modelo es la explicación de enormes beneficios para la interacción social en función de las proximidades del espacio.

En la adopción de estas herramientas el modelo las introduce como supuestos, pero no sin antes relajarlos. Plantea que si la distancia física de los individuos se acerca a cero, los beneficios de la interacción serán elevados pero tendrán un límite. También elabora una especificación del concepto de distancia social que depende tanto de las diferencias en el estatus social logrado entre los individuos, como de las condiciones iniciales captadas por su herencia social; este concepto de distancia social sustituye el cuadrado de la distancia del modelo gravitacional del comercio.

Mediante las especificaciones incluidas, la formulación resultante para el nivel de beneficios esperado por el comercio entre i y j es:

$$\frac{\varepsilon}{(f + d_{o,ij})(g + d_{1,ij}^e)}$$

Donde $d_{o,ij}$ representa la distancia social inicial de los individuos i y j , mientras que $d_{1,ij}^e$, es la distancia final esperada entre los mismos individuos. De la anterior representación se concluye que si

tanto la distancia social inicial, como la distancia social lograda entre dos individuos es cero, las constantes f y g impiden que los beneficios del comercio, sean infinitos. En segundo lugar, un aumento de cualquier distancia entre los individuos, genera un valor de intercambio social asintótico igual a cero.

El modelo está estrechamente relacionado a la primera ley de la geografía, la cual plantea que “(...) todo está relacionado con todo, pero las cosas se relacionan más con las cosas que están cerca que con las cosas que están más lejos” (Tobler, 1970 p.234). La implicación de esta idea de acuerdo con Anselin y Bera (1998) es que el estado de las cosas dependerá en gran parte del estado de las cosas en lugares cercanos. Este proceso de contagio espacial se llama *dependencia espacial*.

La dependencia espacial se explica mediante dos mecanismos, el primero se refiere a que el comportamiento individual se ve afectado por las cercanías espaciales como resultado del proceso de exposición y difusión de acuerdo como lo plantean Morenoff et al (2001) generado por la frecuente socialización de los individuos.

El segundo motivo que justifica el proceso de dependencia espacial son las externalidades espaciales ya sean beneficiosas o perjudiciales que se difunden gracias al efecto de contacto de individuos dentro de la comunidad.

El primer mecanismo provee numerosos aportes teóricos para responder por ejemplo las dinámicas diferenciadas en las comunidades respecto a la supervisión de los adultos sobre miembros más pequeños y la difusión de influencias bien sea positivas o no. También puede explicar otros fenómenos como por qué al vivir en un barrio es más probable encontrar empleo que viviendo en otro. Debido al habitual contacto con individuos que se encuentran laborando y pueden ser fuente para dar información o recomendar en aquellos trabajos.

Mientras que el segundo motivo podría explicar de acuerdo con Flores (2006) por qué un joven con facilidades en el estudio puede ser una fuente de externalidades positivas para los miembros de su clase debido a que el compartir con él aumenta la posibilidad de aprender más fácilmente y por el contrario por qué un niño con comportamientos equivocados es fuente de externalidades negativas en la medida que proporciona información a otros niños que se ubican cerca de él.

El modelo de distancia social junto con el concepto de dependencia espacial proporcionan un aporte teórico inmenso para entender tanto la racionalidad de localización residencial de los individuos, como

la estrecha relación existente en los patrones de comportamiento de cada uno y los factores del entorno, mediante el enfoque de la distancia social que proporciona una visión profunda para examinar los fenómenos económicos y sociales dentro de la estructura general de subculturas o clases.

3. HETEROGENEIDAD EN LAS PREFERENCIAS

En este capítulo se desarrollará a fondo el supuesto de heterogeneidad en las preferencias de los hogares que explica la existencia de subgrupos en la población o entorno social. Es importante tener en cuenta que el entorno social será capturado mediante las dinámicas poblacionales y características de cada barrio.

3.1 COSTOS DE INTERACCIÓN

Es evidente que en el mantenimiento y formación de los lazos sociales se incurren en ciertos tipos de gastos y es por ello que se genera en cierta medida una restricción determinada por la distancia física y factores asociados al transporte. En otras palabras, mantener activos sociales está inversamente relacionado al aumento de la distancia física o tiempo de viaje entre los individuos que residen en distintos barrios; o lo que es lo mismo, el supuesto implica que los agentes son más propensos a interactuar con personas que viven más cerca. En otra medida, las organizaciones dentro de cada barrio juegan un papel importante, instituciones como las iglesias, centros económicos, clubes, escuelas, asociaciones culturales o deportes que se practican en cada uno, son reductores de los costos de las interacciones de los individuos que residen en un barrio o zonas adyacentes, facilitando y favoreciendo de cierto modo el intercambio de información a nivel local y generando un entorno en alguna medida homogéneo para los agentes participantes en esta interacción.

3.2 ENTORNO SOCIAL: RAZA, EDUCACIÓN Y OCUPACIÓN

Como se planteó anteriormente, los agentes tienden a escoger como contactos sociales a personas similares a sí mismos. Se deriva de ahí que las redes personales pueden estar determinadas por atributos socio-demográficos específicos, tales como las características étnico-raciales, religiosas, niveles de educación, sexo, edad, idioma o edad. Estas características propias de cada barrio, constituyen en gran proporción el contenido informativo de las interacciones sociales y la baja calidad de esta información depende en gran parte de la condición de desempleo en cada uno. Desde el punto de vista de este análisis, unos contactos sociales de cada agente son más productivos que otros, dependiendo implícitamente de la capacidad de cada uno para poder transmitir información útil sobre ofertas de trabajo o para poder generar referencias.

La implicación de la utilización de este supuesto sobre este trabajo introduce implícitamente la

hipótesis para probar que existe algún patrón en la distribución espacial de los indicadores del entorno social.

3.3 CLUSTERS ESPACIALES

Según Massey and Denton (1988), el nivel de agrupamiento espacial que presenta un grupo minoritario, conforma una dimensión de la segregación residencial y hace referencia al grado en que las unidades de muestreo en las zonas (barrios en el caso de estudio) habitadas por individuos pertenecientes a minorías colindan entre sí. De esta manera, mayores niveles de cluster implican estructuras residenciales donde las áreas con poblamiento de minorías, están posicionadas geográficamente en vecindades; mientras que menores niveles se caracterizan por configuraciones donde grupos minoritarios están dispersos en el entorno urbano. Los mismos autores identifican varias medidas para determinar el grado de clusters, tal y como lo son el Índice Absoluto de Clusters (ACL), el Índice de Proximidad Espacial (SP) y el Índice Relativo de Clusters (RCL). Para este estudio se empleará el índice de Morán de distribución espacial.

Índice Global de Morán Univariado

Éste índice es una medida de autocorrelación espacial y determina el grado en que cierto fenómeno en un área (barrio) está asociado o se asemeja al mismo en lugares adyacentes. El I de Morán permite rechazar la hipótesis nula de distribución aleatoria de la población según la medida de entorno social que más adelante se proponen en este trabajo. En otras palabras, capta la relación entre cierta variable en un barrio y el promedio de la misma variable en los barrios vecinos. Por ejemplo, si el fenómeno de interés es la población afrocolombiana, el índice muestra el grado en que el porcentaje de población afrocolombiana en un barrio es similar a la de los barrios que lo colindan. La especificación del índice es la siguiente:

$$I = \frac{n \sum_i \sum_j w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{S_0 \sum_i (x_i - \bar{x})^2}$$

$$\text{Y } S_0 = \sum_i \sum_j w_{ij}$$

Donde n, es el número de subunidades geográficas (barrios), x_i y x_j son los valores que toman la variable analizada en el barrio i y j respectivamente. En el caso en que se mira el agrupamiento de grupos afrocolombianos, x_i y x_j representa el porcentaje de individuos con esta característica en los

barrios i y j . Por otro lado, w_{ij} puede tomar valor de uno si los barrios i y j son fronterizos o cero si no lo son, en otras palabras, establece una aproximación al determinar qué pares de barrios se encuentran cerca los unos de los otros y cuáles no. Dado que ciertos barrios pueden tener más vecinos que otros, S_0 es el ponderador de distancia o vecindad.

La insuficiencia de la utilización de este único indicador global de correlación espacial radica en que no identifica áreas donde la agrupación de grupos es significativa. En consecuencia es necesario emplear el índice local de Morán.

Índice Local de Morán Univariado

Corresponde al indicador local de autocorrelación espacial (LISA) más usado. Anselin (1995, p. 94), define estos indicadores LISA estableciendo dos características propias que deben cumplir: La primera es que el LISA para cada observación proporciona una idea de la magnitud de significancia espacial sobre la agrupación de valores similares en torno a la misma; y la segunda es que la sumatoria de las LISAs de las observaciones corresponden a un indicador global de correlación espacial.

La expresión matemática del índice local de Morán es la siguiente:

$$I_i = (x_i - \bar{x}) \sum_j w_{ij} (x_j - \bar{x})$$

El índice local de Morán identifica las áreas donde la agrupación de grupos minoritarios, afrocolombianos por ejemplo, es estadísticamente significativa, estas áreas se denominan “puntos calientes” y corresponderían a barrios con una alta proporción de comunidades afrocolombianas, rodeados por otros con similares características en cuanto a esta observación. De igual forma, el índice muestra “puntos fríos”, que corresponden a barrios con poca proporción de afrocolombianos rodeado de barrios de igual característica; para seguir con el ejemplo.

Este índice univariado de Morán presentado hasta el momento proporciona un gran aporte metodológico para de alguna forma poder cuantificar el grado de dependencia espacial de una variable en una determinada área en relación a zonas vecinas. Sin embargo, en este trabajo también se analiza la correlación espacial existente entre dos variables.

Índice de Morán Multivariado

A fin de encontrar el grado de correlación espacial entre una variable en cierta localización y otra variable en unidades geográficas vecinas, se introduce un coeficiente global de correlación espacial

bivariado representado a continuación:

$$I_{kl} = \frac{Z_k' W Z_l}{Z_k' Z_k}$$

Donde Z_k podría representar el porcentaje de individuos que tienen educación superior o posgrado en cierta unidad espacial, mientras que $W*Z_l$ captura el efecto de la correlación espacial del porcentaje de individuos afrocolombianos en el entorno de dicha unidad, es decir, muestra la relación que existe en cada punto de la muestra, entre los valores de una variable y el valor medio de otra variable en las vecindades de dicho punto.

Para este caso también es acertado incluir índices locales bivariados de correlación espacial. El estadístico LISA tiene en cuenta para cada unidad geográfica, valores de una variable y el respectivo retardo espacial para la otra variable, proporcionando herramientas para determinar áreas de agrupamiento y establecer con significancia estadística distintas asociaciones espaciales entre las dos variables.

4. DATOS

La totalidad de los datos provienen de estadísticas georreferenciadas de los datos censales del Departamento Administrativo Nacional de Estadística (DANE) de 2005 para la ciudad de Cali a nivel de sectores censales, constituyendo un total de 340, base estadística disponible en el Censo Básico. En este trabajo se estudian los patrones que provocan la aglomeración en las tasas de subutilización (denotada por Z10605C) entre los barrios, para explicar este fenómeno se utilizarán los datos correspondientes a:

- : porcentaje de la población en el barrio i mayor de 25 años que tiene nivel educativo superior o posgrado (cuando se hable del indicador de capital humano se hará referencia a esta misma medida)

Donde corresponde al número de personas en el barrio i que tienen más de 25 años y se graduaron de estudios superiores.

Y es el número de personas en el barrio i que tienen más de 25 años.

- : porcentaje de la población afrocolombiana² en el barrio i .
- : Edad promedio de la población en el barrio i .
- : Retardo espacial de la tasa de subutilización para el barrio i ³.

5. ANÁLISIS EXPLORATORIO DE DATOS ESPACIALES

En la primera parte de este capítulo se empleará el Análisis Exploratorio de Datos Espaciales (AEDE) que constituye un campo para el análisis económico y de las ciencias sociales en general. El AEDE

² La población afrocolombiana es el agregado de: mulato, raizal, rom, palenquero, indígena y afrocolombiano. La población no afrocolombiana es el resto de la población.

³ Estos datos fueron obtenidos en el marco del proyecto “Segregación residencial con econometría espacial en Cali” mencionado anteriormente.

podría definirse como el grupo de herramientas estadístico-gráficas que describen y visualizan las distribuciones espaciales, identificando localizaciones atípicas, descubriendo formas de asociación (autocorrelación espacial) que, a su vez, pueden ser de carácter global o local, y sugiriendo estructuras en el espacio geográfico (heterogeneidad espacial). El AEDE puede ser aplicado desde una perspectiva univariante, sobre uno o varios indicadores geográficos, o bivalente, como una etapa propia del proceso econométrico espacial, previa tanto al análisis confirmatorio (estimación y contrastes) como al ejercicio de predicción espacial que será implementado en el próximo capítulo.

Los datos empleados en este análisis serán la edad promedio del total de individuos en cada uno de los barrios, la composición racial y el indicador de capital humano mencionada anteriormente.

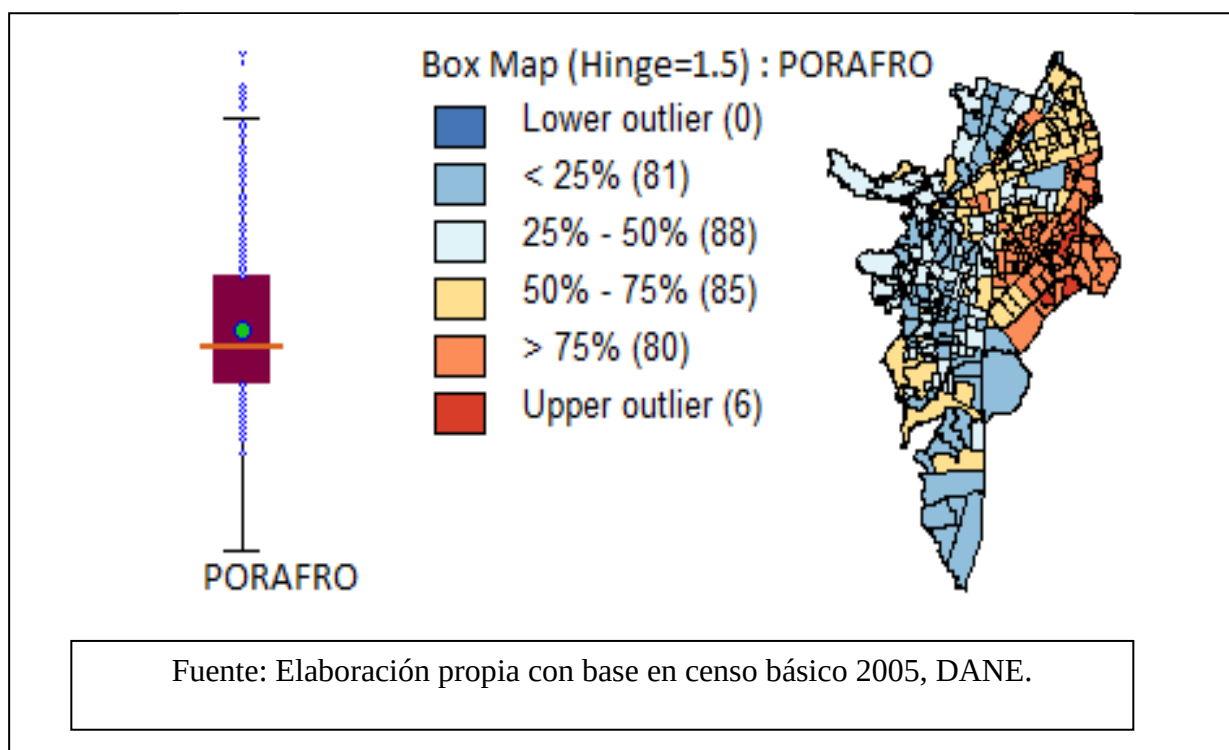
5.1 ANÁLISIS UNIVARIADO

Representación de datos atípicos mediante diagrama/mapa de caja

Los datos atípicos u “outliers” espaciales pueden describirse como valores de la variable que son excepcionalmente bajos o altos, en otras palabras, pueden no ser representativos de la distribución general y afectar el análisis estadístico. Estos pueden ser simplemente errores en la entrada de datos o acontecimientos extraños para los que no existe explicación o bien, representar observaciones interesantes de la población que se pretenden explicar en el estudio. Para el primer caso es común excluirlos para evitar distorsiones en el análisis posterior.

El método del diagrama/mapa de caja está basado en la representación del cálculo tanto de los cuartiles y la mediana de una variable, como de la obtención de las cotas superior e inferior. Se construye una caja de forma que el valor inferior de la misma es el primer cuartil (que contiene el 25% de las observaciones) y el valor superior, el tercer cuartil (que contiene el 75% de las observaciones); mientras que la mediana queda destacada en mitad de la caja con un círculo y una línea horizontal que la atraviesa. Las cotas se obtienen sumando (restando) a la mediana el producto de los valores del tercer (primer) cuartil por 1,5 el recorrido intercuartílico ó 3 veces si se es menos estricto. Los datos atípicos se sitúan por encima (por debajo) de dichas cotas. No necesariamente existen dichos valores, como cuando los datos están concentrados cerca a la media. Por esta misma razón, puede ocurrir que la herramienta no refleje datos atípicos, a pesar de su existencia.

Figura 1: Diagrama de caja y su representación gráfica de la variable porcentaje de afrocolombianos por barrio.



Puede observarse que existen 6 barrios que presentan concentración de población afrocolombiana muy superior al promedio municipal, razón por la que son identificados como outliers. Se distingue también que la zona oriental de la ciudad corresponde al tercer cuartil de la distribución de concentración de afrocolombianos. Cabe aclarar que existen 11 barrios para los cuales no hay datos (que corresponden al punto más bajo del diagrama), pero dadas las características de la población, se encuentran cercanos a los valores de la media y no son considerados como valores atípicos de la variable.

A continuación, dadas las características de la distribución de la variable edad promedio por barrio, se evidencian 17 con valores atípicos. Los 11 que se encuentran considerablemente por debajo de la media, son los mismos para los cuales no se poseen datos, mencionados en la figura 1. Por otro lado, por encima de la cota máxima se encuentran 6 barrios que presentan población con promedio de edad muy superior a la media, localizados en la zona central de la ciudad y son considerados outliers espaciales. También se aprecia una concentración de valores correspondientes al primer cuartil de la distribución, al oriente y occidente del municipio, indicando alta proporción de individuos jóvenes en aquellas áreas de la ciudad.

Figura 2: Diagrama de caja y su representación gráfica de la variable promedio de edad por barrio.

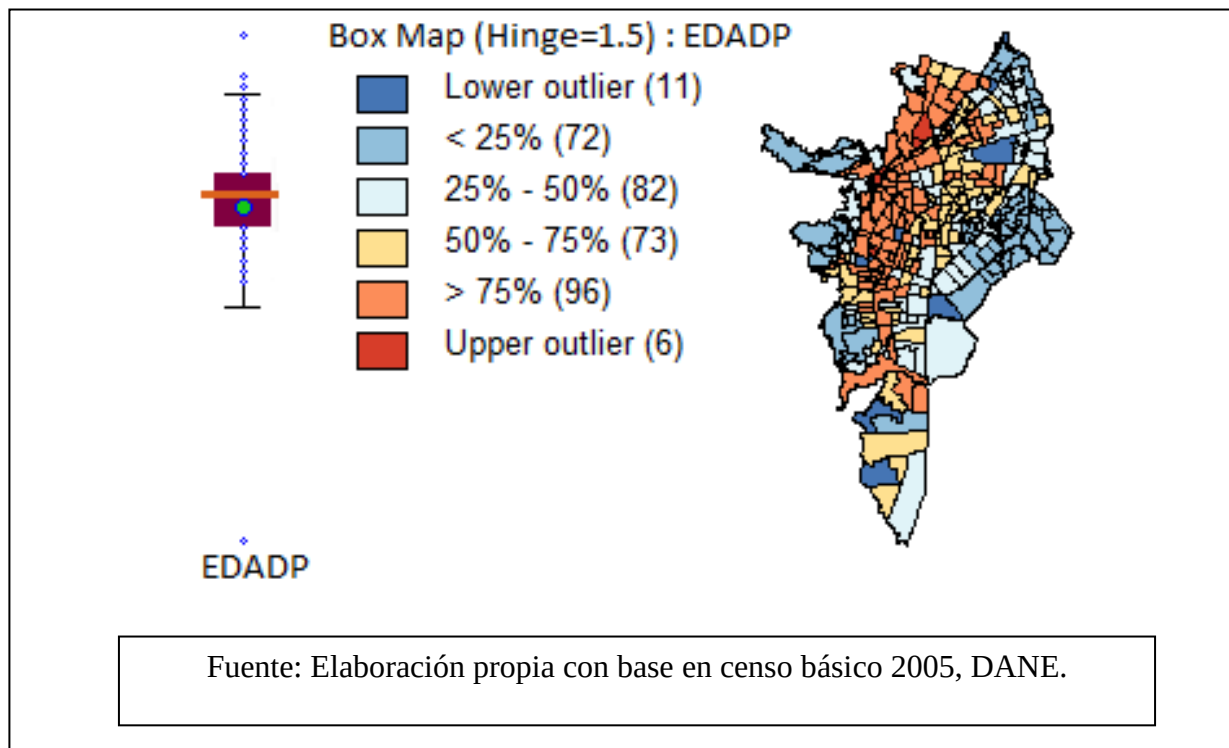
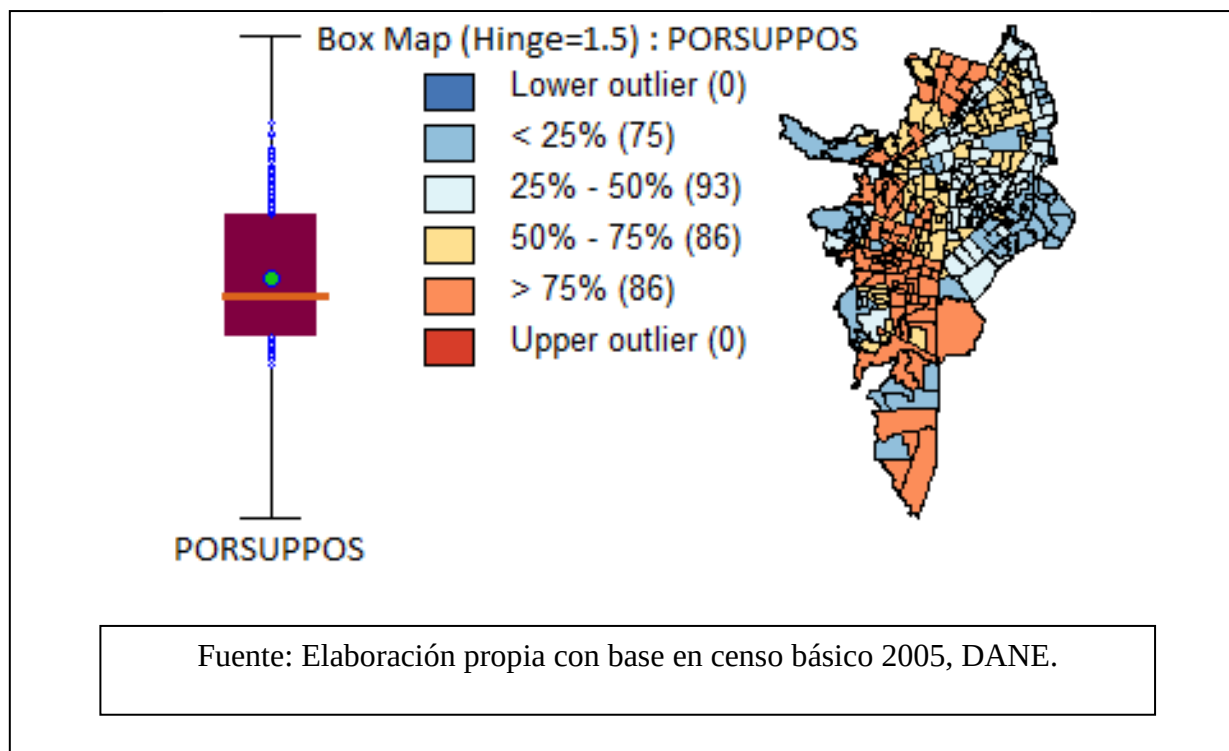


Figura 3: Diagrama de caja y su representación gráfica de la variable porcentaje de personas que tienen nivel académico superior y postgrado.



En este indicador de capital humano, se presentan los peores niveles en toda la zona oriental de la ciudad y en ciertos barrios del occidente. El diagrama no logra reflejar valores atípicos de los 11 barrios para los cuales no se posee información, lo que está explicado por las características de la distribución de la variable. Por otro lado, los máximos nivel de capital humano están ubicados en el eje longitudinal del municipio.

El diagrama de caja permite destacar características propias de las distintas zonas de la ciudad debido a que se aprecian patrones en la distribución de las variables, concentrando valores bajos y altos de las mismas en zonas claramente diferenciadas, lo sugiere la posible existencia de autocorrelación espacial en las variables. Este fenómeno exige el uso de otras herramientas que permitan su medición, como lo son el índice tanto global como local de Morán descrito en el capítulo 4.

Dado que se pretende hacer un análisis local, es preferible para el estudio hacer extrapolación para los 11 barrios al valor promedio de toda la comuna para cada una de las variables, asumiendo que los barrios pertenecientes a una comuna presentan comportamientos similares, esto con el fin de evitar distorsiones en el análisis, puesto que el programa utilizado (geoda) no permite excluir estos valores para análisis locales.

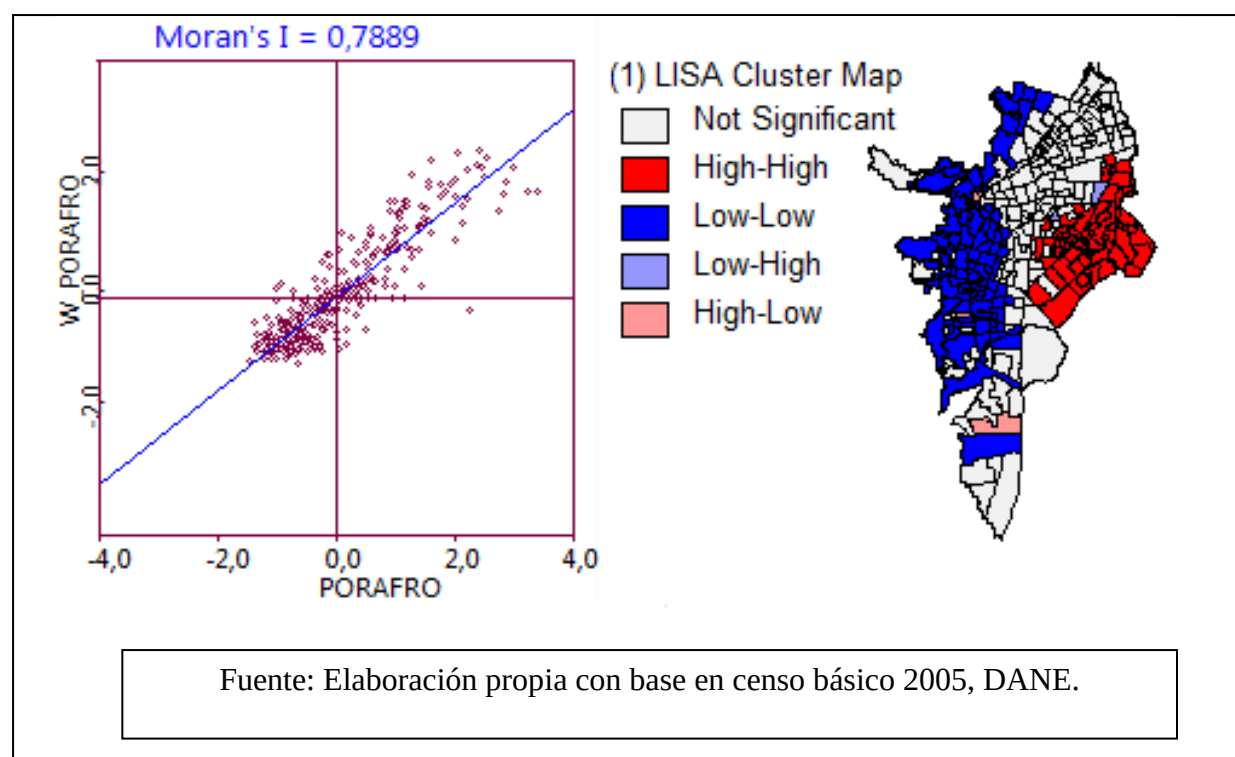
Análisis de autocorrelación espacial

Para establecer si la concentración de la distribución de las anteriores variables es significativa, se medirá el grado de autocorrelación de cada una de ellas en cada barrio, con los valores de las mismas en los barrios vecinos. Este análisis permitirá determinar la existencia o no de cluster espaciales en la ciudad.

Población afrocolombiana

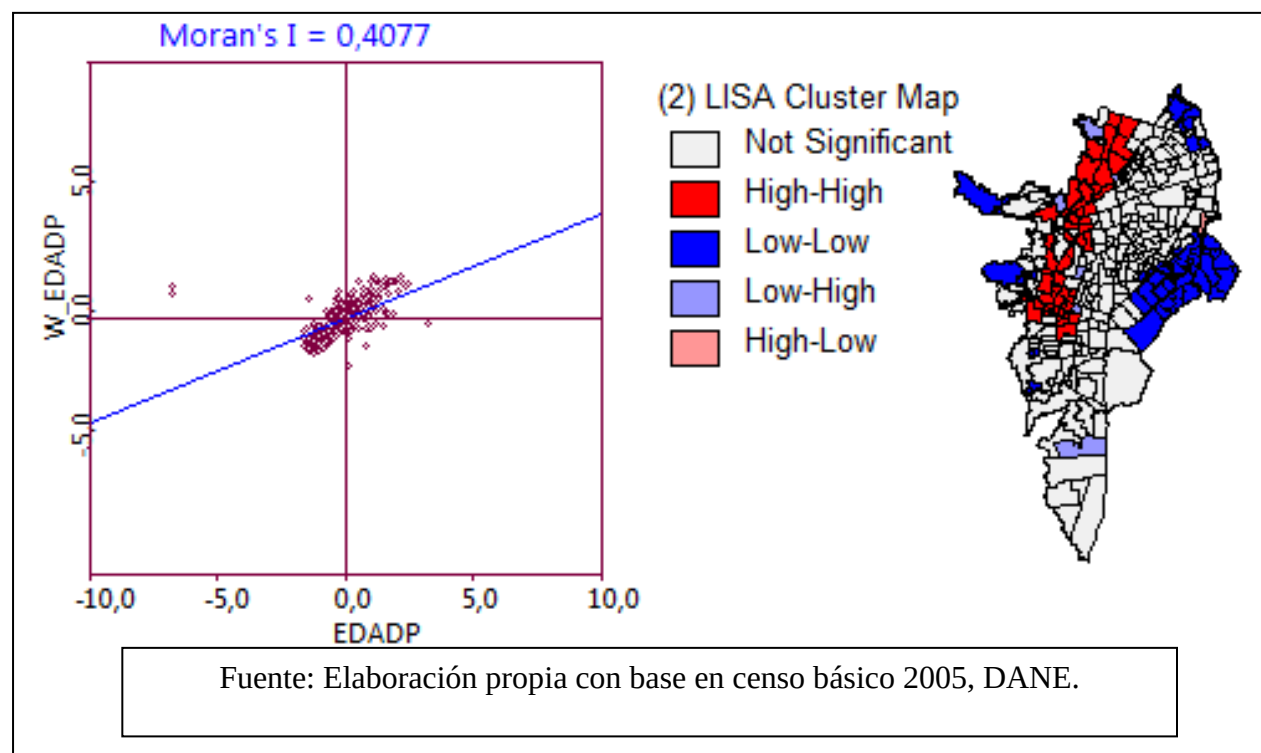
El índice global de Morán de 0.7889 que se observa en la figura 4, refleja la existencia de una alta autocorrelación positiva en la distribución de esta variable entre los barrios. Por un lado, el primer cuadrante indica barrios con alta proporción de población afrocolombiana contiguos a barrios con la misma característica, estos corresponden a los barrios del oriente de la ciudad, señalados con rojo en el mapa. Mientras que los barrios ubicados en el tercer cuadrante, presentan una baja concentración de población afrocolombiana y están rodeados de barrios cuya población es en su mayoría no afrocolombiana; y están marcados de color azul. La alta autocorrelación encontrada, indica la presencia de un cluster de población afrocolombiana y otro de no afrocolombiana; observados en el mapa con el color rojo y azul respectivamente.

Figura 4: Índice Global de Morán y mapa de significancia de clusters para la variable porcentaje de afrocolombianos por barrio.



Edad Promedio

Figura 5: Índice Global de Morán y mapa de significancia de clusters para la variable promedio de edad por barrio.

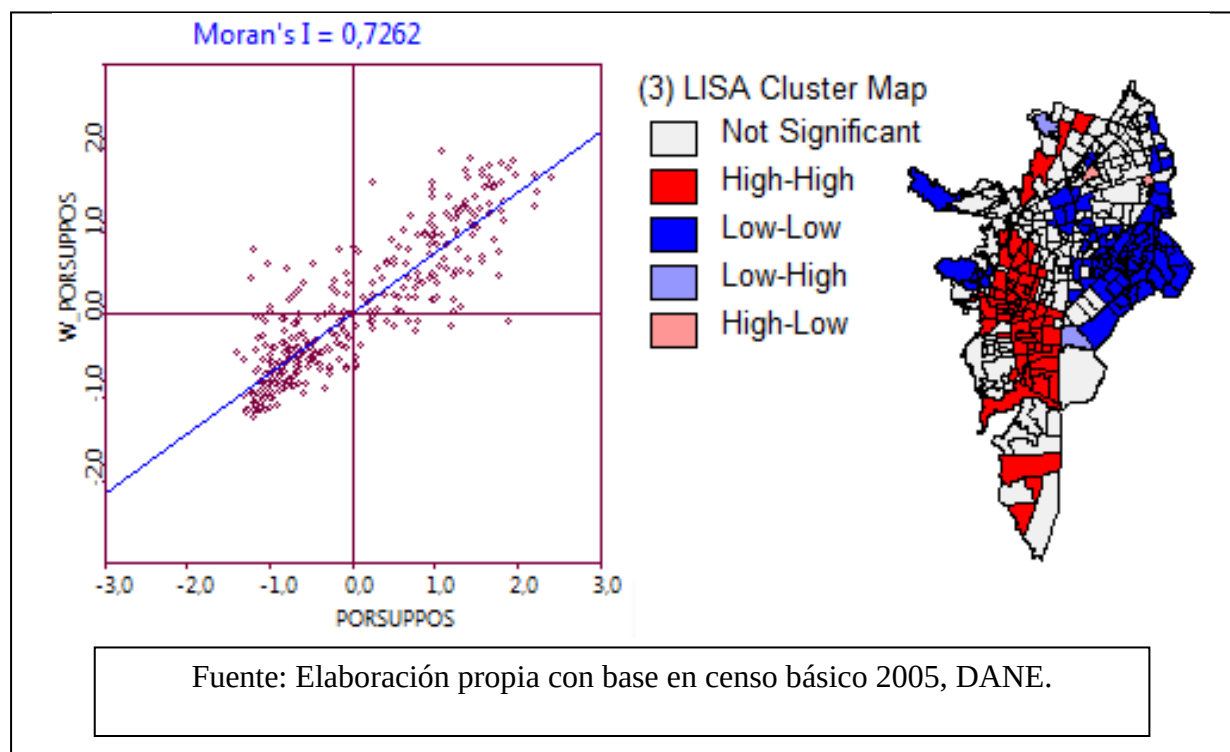


El mapa señala 5 clusters para la variable, 4 de ellos representan barrios con promedios bajos de edad localizados en el tercer cuadrante del índice global de Morán. En el primer cuadrante se sitúan los barrios con mayor promedio de edad, conformando el otro cluster.

Se puede apreciar que la mayoría de los barrios se encuentran en el primer y tercer cuadrante, reflejando la presencia de asociación espacial del índice en cada barrio con sus vecinos.

Población de 25 años o más con educación superior o posgrado

Figura 6: Índice Global de Morán y mapa de significancia de clusters para la variable porcentaje de personas que tienen nivel académico superior o postgrado.



La gran mayoría de los barrios se encuentran en los cuadrantes 1 y 3, lo que explica el alto valor que toma el índice de Morán y sugiere una elevada autocorrelación para la variable. Se evidencia gráficamente en el mapa de cluster grandes aglomeraciones de barrios con baja proporción de individuos que tienen nivel académico superior o posgrado y otras cuantas de barrios con alta proporción de individuos con esta característica.

Observando conjuntamente la distribución geográfica de las tres variables en la ciudad de Cali, se pueden identificar dos grandes áreas. Una localizada al oriente de la ciudad compuesta principalmente por barrios de las comunas 7, 11, 12, 13, 14, 15, 16 y 21, que concentra más del 60% de la población

afrocolombiana de la ciudad⁴, con una estructura poblacional compuesta en su mayoría por jóvenes y baja proporción de personas con estudios avanzados. Y la otra gran área se ubica en todo el cordón central del municipio, que concentra población con características totalmente opuestas a la primera y se compone en su mayoría por barrios de las comunas 2, 3, 9, 10, 17, 19 y 22. Esta fuerte división de la población sugiere la existencia de segregación racial. Para determinar en qué medida este fenómeno está correlacionado con las diferencias encontradas en las otras variables analizadas, es preciso emplear herramientas de análisis multivariado.

5.2 ANÁLISIS BIVARIADO

Análisis de correlación espacial

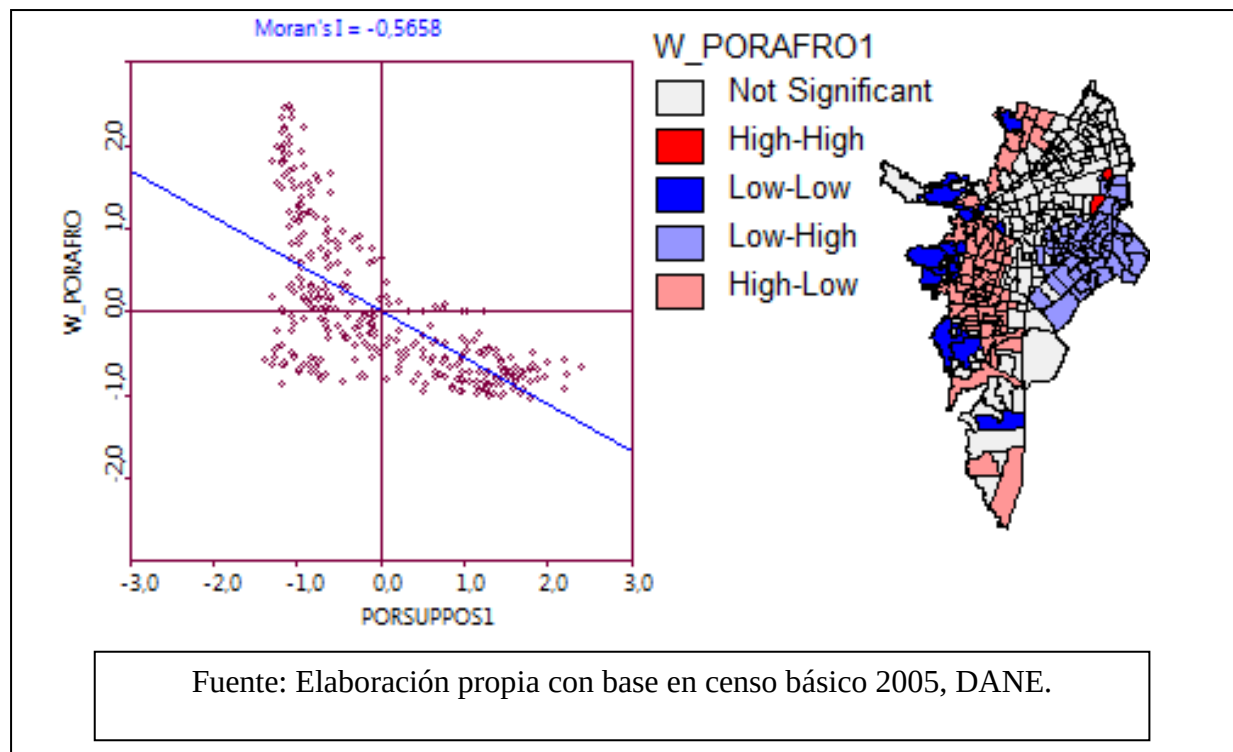
En esta sección se calculará el grado de correlación entre el valor de la variable porcentaje de personas que tienen nivel académico superior y postgrado para cada barrio, con el valor promedio del porcentaje de población afrocolombiana que reside en barrios contiguos y la correlación entre esta última observación y la variable promedio de edad en cada barrio.

Población afro y educación superior y posgrado

En la figura 7 se observan una gran cantidad de barrios representados por puntos en el diagrama que se encuentran en los cuadrantes 2 y 4 indicando una presencia de correlación espacial negativa entre ambas variables evidenciada por el alto índice de Morán de -0,5658. Los barrios reflejados en el cuadrante 2, se caracterizan por presentar bajos valores en el indicador de educación rodeados por barrios con gran proporción de población afrocolombiana. Estos barrios se ubican en el mapa al oriente de la ciudad con azul claro. Por otro lado, los barrios que se encuentran en el cuarto cuadrante son aquellos que presentan altos valores en el indicador de educación y se localizan contiguos a otros cuya población afrocolombiana es reducida, se ubican en el cordón central de la ciudad y están señalados de color rosado. Al occidente de la ciudad se observan 3 pequeñas concentraciones de barrios que presentan bajos valores en el indicador de educación y son vecinos de barrios con baja concentración de afrocolombianos.

⁴ Ver anexo 1

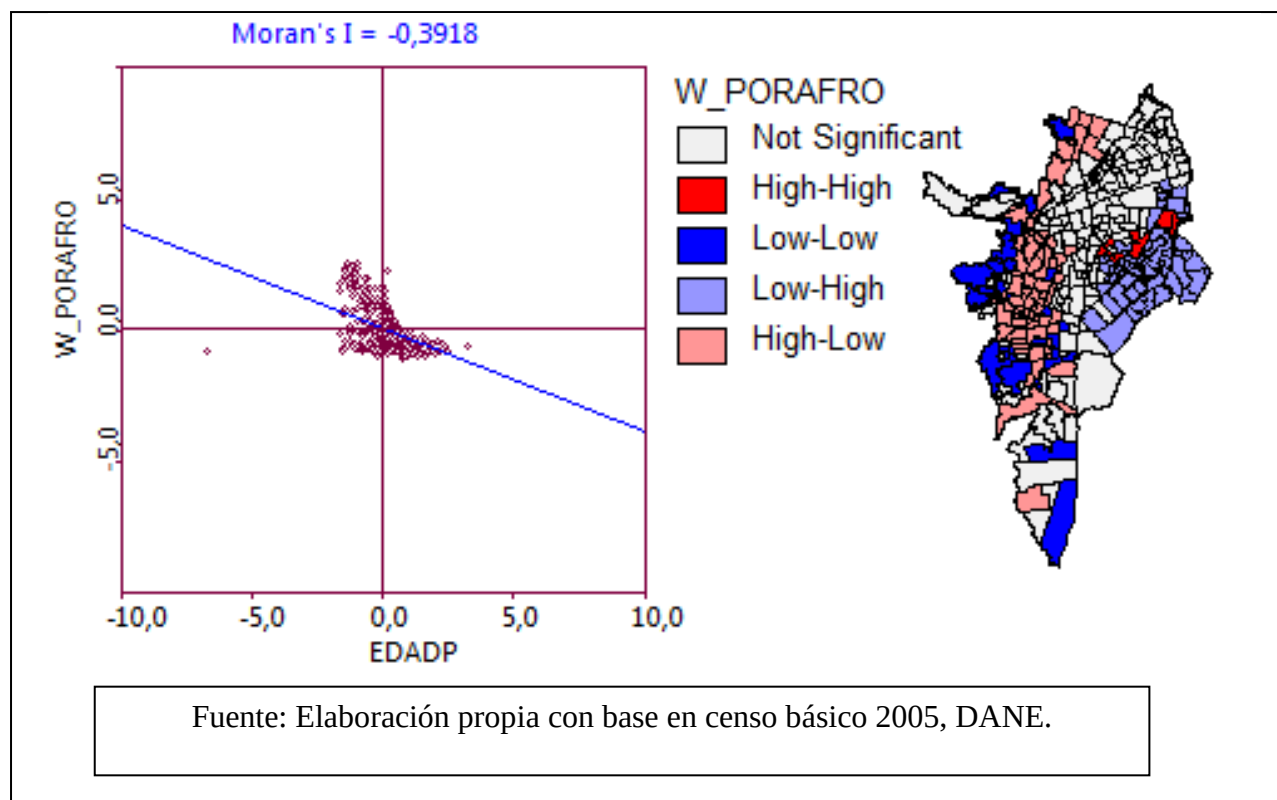
Figura 7: Índice Global bivariado de Morán y mapa de significancia de clusters para las variables porcentaje de población afrocolombiana y porcentaje de personas de 25 años o más que tienen nivel académico superior o postgrado.



Población afro y edad promedio

El índice de Morán de -0.3918 señala la existencia de moderada correlación espacial para estas dos variables, pues a pesar de que la mayoría de los barrios se observan en los cuadrantes 2 y 4 que haría que la correlación fuera fuertemente negativa, existe una considerable concentración de barrios al occidente de la ciudad, donde la correlación es positiva localizados en el cuadrante 3 del diagrama. Señalados de azul claro en el mapa, se ubican los barrios del segundo cuadrante del diagrama que presentan bajo promedio de edad y están rodeados por barrios con alta proporción de población afrocolombiana. De rosado se encuentran los barrios del cuadrante 4 y se caracterizan por alto promedio de edad y baja proporción de afrocolombianos en los barrios vecinos. Y de azul oscuro por barrios del occidente de la ciudad y su población, en su mayoría jóvenes, y cuyos vecinos tienen bajas tasas de población afrocolombiana. La figura 8 a continuación evidencia las características mencionadas.

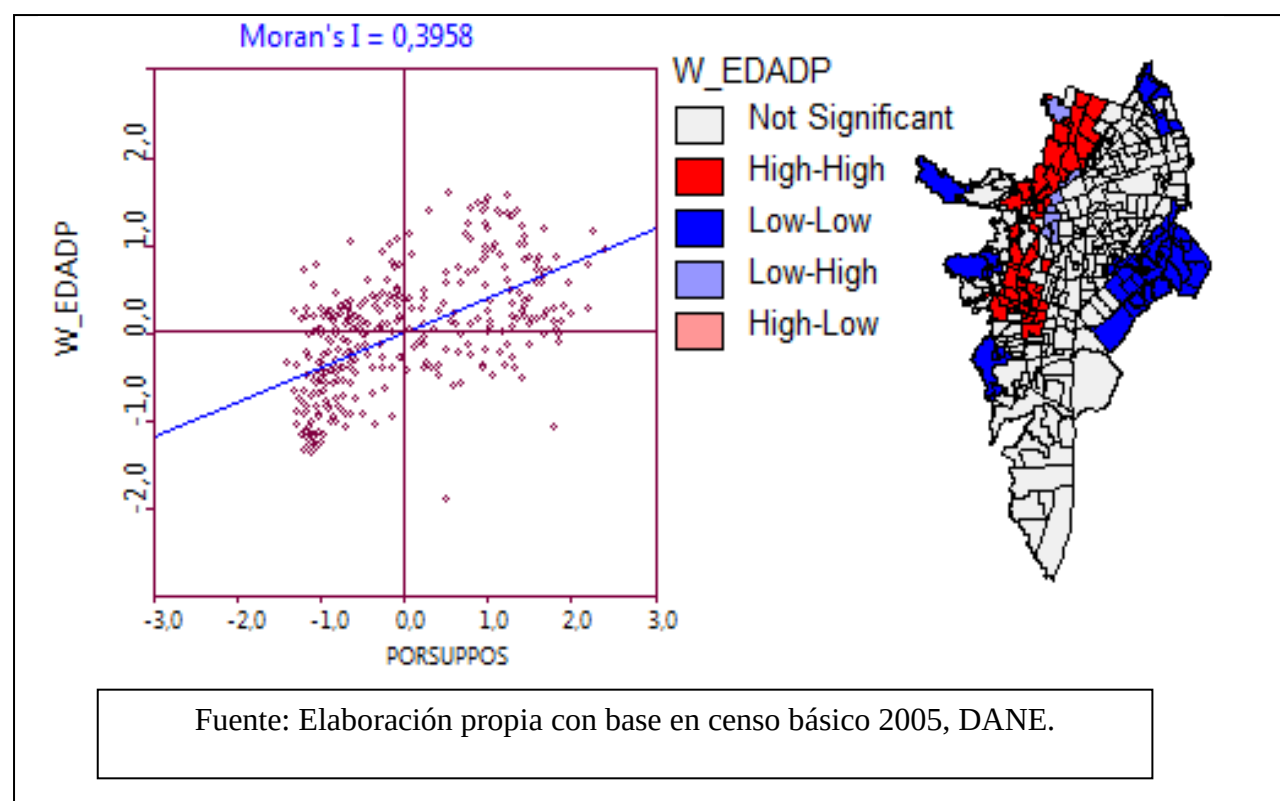
Figura 8: Índice Global bivariado de Morán y mapa de significancia de clusters para las variables porcentaje de población afrocolombiana y edad promedio por barrio.



Edad promedio y capital humano

Se evidencian en la figura 9 claramente 2 tipos de clusters principalmente para las dos variables, por un lado y localizados al oriente y occidente de la ciudad se encuentran barrios caracterizados por tener bajos niveles de capital humano y estar rodeados por otros que presentan población con bajo promedio de edad. Estos barrios se encuentran en el tercer cuadrante del índice globalizado de Morán. Por otro lado, los barrios observados de color rojo son aquellos que sus vecinos más cercanos presentan una población con mayores promedios en la edad y en ellos existe altos indicadores de capital humano.

Figura 9: Índice Global bivariado de Morán y mapa de significancia de clusters para las variables porcentaje de personas de 25 años o más que tienen educación superior o posgrado y edad promedio en los barrios.



6. ANÁLISIS CONFIRMATORIO DE DATOS ESPACIALES

Test de autocorrelación global

Los estadísticos globales de autocorrelación constituyen la aproximación más tradicional al efecto de dependencia espacial, permitiendo contrastar la presencia o ausencia de un esquema de dependencia espacial a nivel univariante, es decir, contrastar si se cumple la hipótesis de que una variable se encuentra distribuida de forma totalmente aleatoria en el espacio o si por el contrario, existe una asociación significativa de valores similares o disímiles entre regiones vecinas.

Para realizar el contraste de hipótesis de no existencia de autocorrelación espacial, se utiliza el estadístico de I- Moran.

Tabla 1: Medida de autocorrelación global I de Morán de la variable de capital humano estimado

Variable	I de Morán	p-valor
EDUESTIMAD	10,91	0,000

Fuente: Elaboración propia con base en censo básico 2005, DANE

El test de Moran (I) establece como hipótesis nula la no correlación espacial, por lo que un valor significativo positivo (negativo) del estadístico estandarizado $Z(I)$ informa sobre la presencia de un esquema de autocorrelación espacial positiva (negativa), es decir, la presencia de valores similares (disímiles) de la variable entre barrios vecinos.

Teniendo en cuenta lo anterior y mediante la estimación realizada del test, se puede afirmar que la variable de capital humano estimada (EDUESTIMAD), presenta autocorrelación espacial positiva, lo que implica valores similares alrededor de los barrios.

6.1 MODELO DE REGRESIÓN ESPACIAL CON EXTERNALIDADES GLOBALES

Las externalidades espaciales son implementadas en el campo de la economía por medio de los modelos de interacción espacial que introducen interdependencia entre los agentes. Las externalidades globales se modelizan en este estudio por medio de la interacción espacial que surge del efecto de vecindades inter-barrios. Tal y como lo plantea Chasco (2003) la comprobación de la existencia y la medición de la fuerza de estas externalidades espaciales requieren de especificación y estimación de

modelos empleando econometría espacial.

La autora parte de la forma reducida del modelo de externalidades planteado por Anselin (2001), que capta estas externalidades o “spillovers” a través de multiplicadores espaciales y originan numerosas especificaciones dependiendo de los retardos espaciales que afectan a las variables ya sea la dependiente (Wy), las explicativas (WX) y/o a la perturbación aleatoria (Wu).

Test de diagnóstico de identificación del modelo

Una vez evidenciada la existencia de autocorrelación espacial a partir del test global de Morán, el siguiente paso es elegir el modelo espacial más adecuado, ya sea una estructura que capte el efecto del retardo espacial de la variable dependiente (modelo de rezago espacial) o bien, el retardo espacial en la perturbación aleatoria (modelo de error espacial). Para ello se aplican los test de multiplicadores de lagrange (LM) y sus variantes robustas para la variable estimada de capital humano.

Tabla 2: Test LM y LM robusto de EDUESTIMAD

Test	Estadístico
LM (lag)	110,88*
Robust LM (lag)	2,68***
LM (error)	112,33*
Robust LM (error)	4,13**

* Parámetros significativos al 1%, ** significativos al 5%, *** significativos al 10%.

Fuente: Elaboración propia con base en censo básico 2005, DANE

Tanto el estadístico LM sobre el retardo espacial de la variable endógena, como el LM sobre el retardo espacial de la perturbación aleatoria resultan significativos, al igual que sus respectivos test de robustez aunque a diferente significancia. Debido a ello, no hay suficiente evidencia aun para determinar qué modelo es superior al otro por lo que se optará por hacer la estimación de ambos y mediante otros criterios hacer la elección para posteriormente comentar los resultados del modelo escogido.

Estimación por el método de variables instrumentales (VI)

Debido a la fuerte relación entre capital humano, población afrocolombiana y edad promedio evidenciado en el análisis descriptivo, podría explicarse por qué barrios que tienen alta proporción de población afrocolombiana y una baja edad promedio, presentan bajos niveles de capital humano.

Esta relación podría provenir de las características propias de la población afrocolombiana caleña, en su mayoría inmigrantes de la costa pacífica colombiana, zona que concentra bajos niveles de capital social. A su vez la baja edad promedio puede explicar bajos niveles de capital humano debido a que un barrio con alta proporción de población joven disminuye la probabilidad de que las personas tengan un acompañamiento adecuado desde su niñez generando una menor acumulación de capital social.

Debido a esta endogeneidad de la variable de capital humano, se empleará el método de variables instrumentales.

La implementación de éste método constituye una alternativa más robusta para la estimación de modelos con la variable dependiente espacialmente retardada, pues no requiere la hipótesis de normalidad en la distribución de la perturbación aleatoria. En concreto, se estimará un modelo de mínimos cuadrados espaciales en dos etapas.

La estimación de variables instrumentales está basada en la existencia de un conjunto de elementos que están fuertemente correlacionados con una o más variables explicativas, y asintóticamente incorrelacionados con la perturbación aleatoria. Después de encontrar estos instrumentos, se emplearán en la construcción de una variable estimada de la(s) variable(s) endógena(s) para posteriormente a partir de sus valores estimados mediante mínimos cuadrados ordinarios (MCO), incluirlos en la regresión junto con las variables exógenas del modelo si ellas existen; además de la variable dependiente espacialmente retardada Wy , para el caso del modelo de rezago espacial y el retardo espacial de la perturbación aleatoria Wu para el modelo de error espacial; que se introducen como una variable explicativa adicional de cada modelo y se estima por el método de máxima verosimilitud (MV).

La forma general del modelo en dos etapas que se empleará es el siguiente:

Donde

- Y , es la variable dependiente.
- Z , es el retardo espacial de la variable dependiente y la matriz X de regresores.
- Q , el conjunto de instrumentos correlacionados fuertemente con las variables Z .

La forma estructural del modelo de rezago espacial será:

(1)

(2)

Mientras que para el modelo de error espacial será:

(1)

(2)

6.2 ESTIMACIONES

A continuación se muestra la estimación mediante mínimos cuadrados ordinarios de la variable endógena PORSUPOS empleando como instrumentos PORAFRO y EDADP.

Tabla 3: Estimaciones primera etapa por MCO

	MCO
CONSTANTE	21,2*
EDADP	1,14*
PORAFRO	1,08*
R^2 ajustado	0,5

*Parámetros significativos al 1%

Fuente: Elaboración propia con base en censo básico 2005, DANE

La estimación de la variable porcentaje de personas de 25 años o más con educación superior o posgrado, confirman y cuantifican el impacto de dos fenómenos interesantes como lo son la edad promedio en cada barrio y su composición racial.

Por un lado, el coeficiente del porcentaje de población afro en cada barrio (PORAFRO) de -1,077, indica que cuando aumenta un 1% la población afrocolombiana en un barrio, el indicador de capital humano disminuiría poco más que proporcional, o sea, más de un 1%. Este impacto negativo de la variable puede explicarse dada la naturaleza de los barrios que integran mayor concentración de población afrocolombiana, barrios marginados de la ciudad de Cali ubicados al oriente de la ciudad,

sector conformado por las comunas 13, 14, 15 y 21 principalmente que conforman el llamado distrito de aguablanca, el cual presenta altas tasas de inmigrantes provenientes de zonas rurales del departamento de Chocó y de la ciudad de Tumaco en su mayoría como lo señalan Barbary y Hoffman (2004), además de los barrios que conforman las comunas 7,8, 11 y 12, ubicados también en el extremo este de la ciudad. Barbary (2004) destaca que la evidente segregación de la población afrocolombiana tiene repercusiones importantes sobre el menor acceso a los servicios públicos y en los indicadores de hacinamiento, fenómeno que se corrobora en las “Recomendaciones para complementar el diagnóstico del plan de desarrollo municipal de Cali 2008-2011”⁵ donde se destaca que las comunas con mayor participación de afrocolombianos, presentan mayores tamaños en los hogares y mayores porcentajes de inasistencia escolar.

La relación encontrada entre capital humano y población afrocolombiana evidencia el bajo capital social de la población en estas comunas y los entornos desfavorables en que crecen los niños que no incentivan la acumulación de capital humano para los jóvenes, debido a que no existe en sus barrios medios de desviar la trampa de pobreza en la que se encuentran.

La otra variable de la regresión, edad promedio en cada barrio también tiene un alto efecto sobre el capital humano, el coeficiente EDADP indica que con aumento de 1 año en esta variable, el porcentaje de personas con educación superior o posgrado crece un 1,146. Este estimador puede estar captando los barrios con menor proporción de adultos que estén en contacto con los jóvenes y niños y en consecuencia estos tienen menor probabilidad de estar bajo la supervisión de un grupo de influencias supuestamente positivas, como lo deben ser sus padres y familiares. Esta menor supervisión, tal y como lo plantea Ainsworth (2002) puede también indicar que los niños y jóvenes, sean más propensos a desarrollar actitudes y comportamientos no apropiadas para alcanzar altos niveles de capital humano, además de provocar el aumento del abandono de la escuela.

Otro mecanismo de transmisión que puede estar afectando los indicadores de capital humano a través de la edad promedio de los habitantes de cada barrio, es la calidad del entorno de cada uno. Tal y como se evidenciaba en la distribución porcentual de la edad promedio empleada antes, la mayor concentración de población joven en los barrios se encuentra en la periferia de la ciudad, esto es, el extremo este y oeste.

De cierto modo, las características propias de un barrio como lo son el porcentaje de población

⁵ Ver Anexo 2

afrocolombiana y la edad promedio, captan la calidad del entorno en cada barrio y repercute fuertemente sobre la acumulación de capital humano. Esta estrecha relación entre estatus socioeconómico y nivel educativo puede también verse reflejada por el hecho de que las personas con menores capacidades económicas cuentan con menores recursos educativos (libros, computadores). Esta relación también se puede ver en el sentido inverso, debido a que menores recursos educativos generan menor posibilidad de inversión en capital humano ocasionando una menor productividad que imposibilita conseguir empleos mejor remunerados, lo que está directamente relacionado con el nivel socioeconómico. Esta asociación entre bajo nivel socioeconómico, menores recursos educativos y pobre acumulación de capital humano representa un círculo vicioso.

A continuación se incluyen los valores estimados de la variable endógena en la ecuación (1) para la estimación por medio de ML.

Tabla 4: Estimaciones segunda etapa

	MV (lag)	MV (error)
VARIABLE		
W_Z10605C (Rho)	0.5964*	-----
W_u (Lambda)	-----	0,6064*
CONSTANTE	2,53*	5,43*
EDUESTIMAD	-0,025*	-0,037*
Ln L	-632,31	-634,44
R^2 ajustado	0,43	0,42

*Variables significativas al 1%

Fuente: Elaboración propia con base en censo básico 2005, DANE

Tanto el coeficiente autorregresivo espacial de primer orden correspondiente a la perturbación aleatoria (Lambda), como el rezago espacial de la variable dependiente (Rho), son significativos al 1%. Por otro lado, tanto el logaritmo de la verosimilitud (Ln L) de -632,31, como el R^2 de 0,43 del modelo de rezago espacial, superan los valores obtenidos en el modelo de error espacial de -634,44 y 0,42 respectivamente.

Cabe resaltar que también se realizó la estimación de esta segunda etapa por MCO confirmándose también su inferioridad en relación a ambos modelos espaciales⁶.

La estimación por medio de rezago espacial de la variable porcentaje de subutilización por barrio capta el impacto de la misma variable en barrios vecinos y el efecto del capital humano en cada barrio.

Por un lado, dado un aumento del indicador estimado de capital humano en cada barrio, la tasa de subutilización presenta disminuciones significativas al 1%, lo que indica que en los barrios de Cali el factor “nivel educativo” influye en cierta medida que las personas estén empleadas o no.

Tal y como se mencionó anteriormente, los bajos niveles de capital humano se concentran principalmente en los barrios marginados del oriente y occidente de Cali, que presentan además los más bajos estatus socioeconómicos y menor capital social de la ciudad. Este factor social es un importante elemento que impide el apoyo y transmisión de grandes aspiraciones educativas a los hijos, por lo que difícilmente se logrará mitigar el efecto que tiene el factor económico sobre la calidad y cantidad de capital humano invertido.

Por otro lado, los resultados obtenidos para este modelo confirman que tiene sentido incluir un retardo espacial instantáneo de la variable dependiente ($W_Z10605C$) como un factor más explicativo. El valor de $RHO=0,596$ y significativo. El signo del parámetro indica que si un barrio tiene como vecinos otros barrios con alto índice de subutilización, es muy probable que el mismo presente esta misma característica; dicho de otro modo, existen áreas de concentración tanto de altas como de bajas tasas de subutilización en la ciudad de Cali.

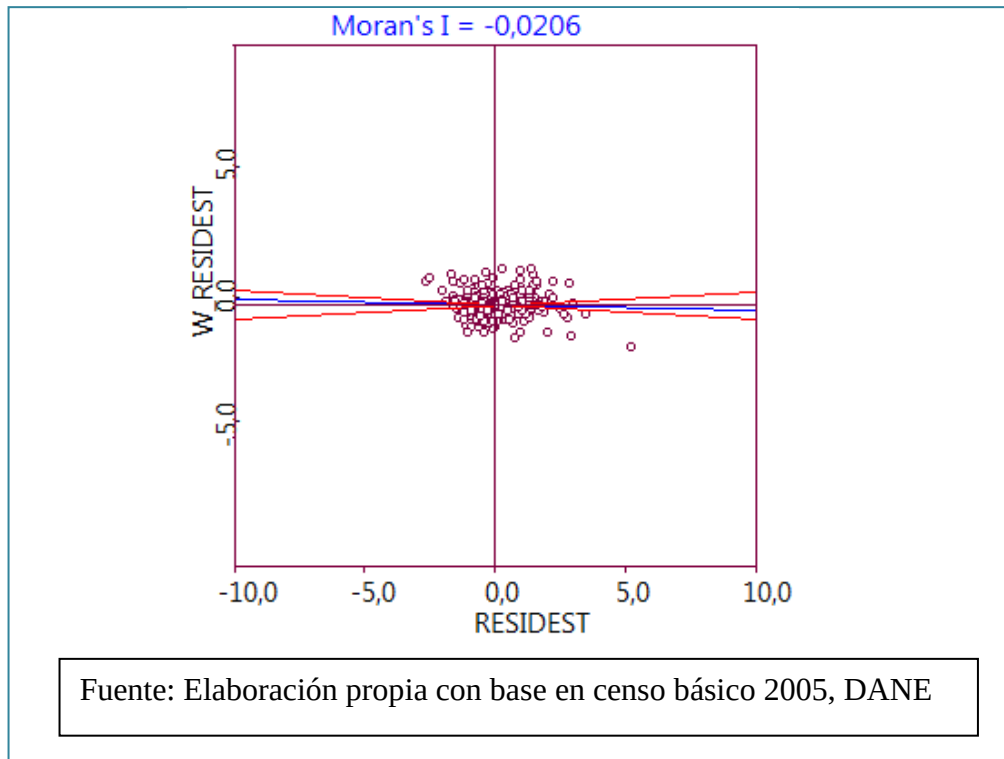
Por último, para evaluar la consistencia y validación del modelo, se presenta a continuación el diagrama de dispersión de Morán para los errores determinados por la diferencia entre los valores observados de la tasa de subutilización por barrios y los valores estimados de la misma.

En la figura 10 se observa que la pendiente tiende a cero y está dentro de la región de aceptación de la hipótesis nula de no autocorrelación local en los residuales del modelo (como se aprecia en la figura 11), lo que indica que la inclusión de la variable dependiente espacialmente rezagada elimina de manera satisfactoria el efecto de segundo orden dentro del modelo de regresión. Esta aseveración se puede ratificar mediante la aplicación del histograma de los resultados obtenidos, a través del test de

⁶ El logaritmo de la verosimilitud arrojado por el modelo en dos etapas por MCO, es de -675,88 que no supera a los obtenidos por los modelos de retardo y error espacial, confirmándose nuevamente su inferioridad de predicción respecto a ellos.

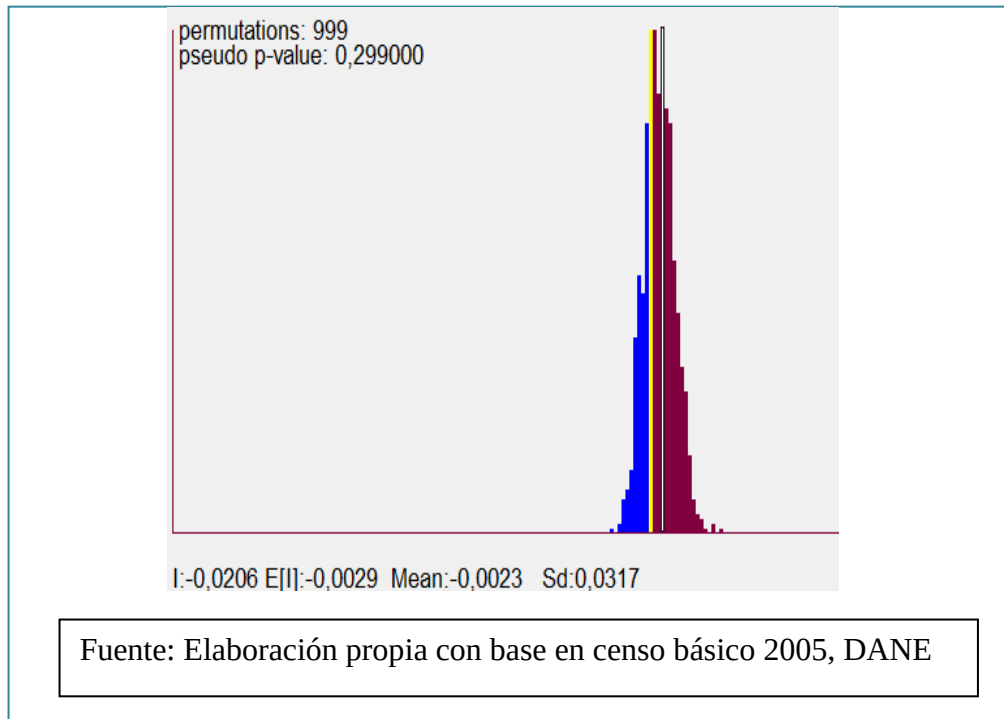
permutaciones.

Figura 10: Dispersión de Morán para los residuales del modelo de rezago espacial.



El histograma que se muestra a continuación corresponde a la aplicación de la I de Moran de los residuales del modelo de rezago espacial para los 340 sectores censales de la ciudad de Cali, a través de un análisis de 999 permutaciones. La I de Morán obtenida es de -0,0206 y se encuentra muy cercana al valor esperado para ese índice que indica que la configuración espacial de los errores es aleatoria $E(I) = -0.0029$, con un p-valor de 0,299. Cae en la región de aceptación de la hipótesis nula (p-valor > 5%) de no autocorrelación espacial en los residuales.

Figura 11: Test de permutaciones



7. CONCLUSIONES

En este trabajo se han estudiado las características de los patrones espaciales en Cali y el efecto que tienen sobre el nivel de capital humano en cada barrio de la ciudad e indirectamente sobre una aproximación a la tasa de desempleo en cada uno.

El análisis exploratorio de datos revela marcadas concentraciones de ciertas características sociales en determinadas zonas de la ciudad que sugieren la existencia de clústeres espaciales de poblaciones minoritarias que están en desventaja frente al resto de la población. Las particularidades encontradas en este grupo poblacional son:

- Alta proporción de afrocolombianos con población joven al oriente de la ciudad y baja proporción de afrocolombianos con población joven al occidente.
- Bajos niveles de capital humano y altas tasas de subutilización en las mismas zonas afectando tanto a niños como jóvenes, adultos y ancianos que gozan de menor acceso a servicios públicos, menores recursos alimenticios y escolares generando menores oportunidades de alcanzar un alto nivel educativo y dificultando el acceso al mercado laboral y más aun a un buen puesto de trabajo.
- La población que reside en estas zonas cuenta con mínimas posibilidades de contacto con influencias positivas que representen un gran potencial de acercamiento al mercado laboral o a puestos de trabajo bien remunerados lo que dificulta revertir la trampa de pobreza en la que se encuentran.

En el análisis confirmatorio se cuantificaron los impactos de las variables utilizadas como aproximación del entorno social sobre el capital humano, como lo son la proporción de población afro y edad promedio en cada barrio, donde cada una recoge el signo esperado para la ciudad de Cali y encontrándose:

- A mayor proporción de afrocolombianos disminuye el indicador de capital humano.
- A mayor proporción de jóvenes en un barrio disminuyen los niveles de educación para las personas de 25 años y más.

Posteriormente se midió el efecto del nivel de capital humano estimado para la población de 25 años y más para cada barrio junto a los efectos de vecindad de la tasa de subutilización sobre la misma tasa en cada barrio, obteniendo los resultados que se esperaban:

- Es muy fuerte el efecto vecindad, o sea que un barrio que tiene como vecinos a otros con altas tasas de subutilización, absorbe en gran parte estas externalidades negativas.
- Y a medida que disminuye el capital humano en cada barrio aumenta la tasa de subutilización en el mismo.

Es preciso señalar que el limitado acceso a datos sobre los perfiles y categorías ocupacionales, en lugar de la tasa de subutilización, así como información sobre precio promedio de la vivienda en cada barrio, constituye una restricción de los alcances de este estudio. No obstante, esta es una línea de gran interés para futuras investigaciones

REFERENCIAS BIBLIOGRÁFICAS

- Ainsworth, James (2002). “Why does it take a village? the mediation of neighborhood effects on educational achievement”. *Social Forces*, Septiembre de 2002, pp.117-152.
- Akerlof, George (1997). “Social distance and social decisions”. *Econométrica*, Vol. 65, No. 5 (Sep., 1997), pp. 1005-1027.
- Anselin, Luc (1995). “Local indicators of spatial association-LISA”. *Geographical Analysis*, 27, pp. 93-115.
- _____.(2001). “*Spatial externalities, spatial multipliers and spatial econometrics*” Discussion Paper del Regional Economics Applications Laboratory, REAL 01-T-11.
- Anselin, Luc y Bera, Anil (1998). “Spatial dependence in linear regression models with an introduction to spatial econometrics”. In: A. Ullah and D. E. A. Giles, Eds., *Handbook of Applied Economic Statistics*, pp. 237–289. New York: Marcel Dekker.
- Barbary, Olivier (2004). “El componente socio-racial de la segregación en Cali”, En: Barbary, Olivier & Urrea, Fernando. *Gente negra en Colombia, Lealon*, Medellín, Marzo de 2004, pp, 157-194.
- Barbary, Olivier & Hoffman, Odile (2004) “La costa pacífica y Cali, sistema de lugares”, En: Barbary, O. & Urrea, F. *Gente negra en Colombia, Lealon*, Medellín, Marzo de 2004, pp. 113-153.
- Bayer, Patrick, McMillan, Robert. and Rueben, Kim (2001). “The causes and consequences of residential segregation: an equilibrium analysis of neighborhood sorting”
- Borjas, George (1995). “Ethnicity, neighborhoods, and human-capital externalities”. *The American Economic Review*, Vol. 85, No. 3 (Jun., 1995), pp. 365-390
- Casas, Andrés, Gamboa, Luis y Piñeros, Luis (2002) “El efecto escuela en Colombia 1999-2000” Serie documentos. *Borradores de Investigación*. Universidad del Rosario. Setiembre 2002.
- Chasco, Coro (2003) “Econometría espacial aplicada a la predicción-extrapolación de datos microterritoriales”. Consejería de Economía e Innovación Tecnológica, Comunidad de Madrid. Madrid.

- Conley, Timothy and Topa, Giorgio (2002). "Socio-economic distance and spatial patterns in unemployment". *Journal of applied econometrics*, Vol. 17, No. 4 (Jul. – Ago. 2002), pp. 303-327.
- Considine, Gillian y Zappalà, Gianni (2002). "Factors influencing the educational performance of students from disadvantaged backgrounds", en T. Eardley and B. Bradbury, eds, *Competing Visions: Refereed Proceedings of the National Social Policy Conference 2001, SPRC Report 1/02*, Social Policy Research Centre, University of New South Wales, Sydney, 91-107.
- Cutler, David and Glaeser, Edward (1997). "Are ghettos good or bad?". *The Quarterly Journal of Economics*. Vol. 112, No. 3 (Aug., 1997), pp. 827-872.
- Duncan, Otis and Duncan, Beverly (1955). "Residential distribution and occupational stratification". *The American Journal of sociology*, Vol. 60, No. 5, World urbanism (Ago., 1955), pp. 493-503.
- Flores, Carolina (2003). "Consecuencias de la segregación Residencial. Teoría y métodos". Pontificia Universidad Católica de Chile.
- _____.(2008). "Residential segregation and the geography of opportunities: a spatial analysis of heterogeneity and spillovers in education". Tesis doctoral, The University of Texas at Austin.
- Forero, Germán (2007). "Efectos del capital étnico y de vecindad sobre la acumulación de capital humano: el caso de las poblaciones de ascendencia afrocolombianas de Cali". Tesis de maestría en economía, Universidad de la Salle.
- Kain, John (1968). "Housing Segregation, negro employment, and metropolitan decentralization." *Quarterly Journal of Economics* 82:175–197.
- Massey, Douglas and Denton, Nancy (1988). "The Dimensions of residential segregation". *Social Forces*, Vol. 67, No.2 (Dic., 1988), pp. 281–315.
- Morenoff, Jeffrey; Sampson, Robert y Raudenbush, Stephen (2000). "Neighborhood inequality, collective efficacy, and the spatial dynamics of urban violence". *Criminology* 39, 517-59.
- Tobler, Waldo (1970). "A computer movie simulating urban growth in the Detroit región". *Economic Geography* 46, 234-40.
- Topa, Giorgio (2001). "Social interactions, local spillovers and unemployment". *The Review of Economic Studies*, Vol.68, No.2 (Apr., 2001), pp. 261-295.
- Vivas, Harvy (2008) "Educación, background familiar y calidad de los entornos locales en

Colombia” Tesis doctoral. Universidad Autónoma de Barcelona.

- Wheeler, James and Stutz, Frederick (1971). “Spatial dimensions an urban social travel”. *Annals of the Association of American Geographers*, Vol. 61, No. 2 pp. 371-386.

ANEXOS

Anexo A

Comuna	Población	Porcentaje de Afrocolombianos	Total Afrocolombianos
Comuna 1	61999	11.45	7099
Comuna 2	102080	10.22	10433
Comuna 3	44308	15.1	6691
Comuna 4	55965	22.92	12827
Comuna 5	99844	20.86	20827
Comuna 6	169392	20.44	34624
Comuna 7	77775	37.54	29197
Comuna 8	97335	19.13	18620
Comuna 9	48382	16.55	8007
Comuna 10	103671	15.34	15903
Comuna 11	98361	33.27	32725
Comuna 12	67439	37.35	25188
Comuna 13	169829	37.82	64229
Comuna 14	151789	51.04	77473
Comuna 15	126709	49.81	63114
Comuna 16	94445	26.86	25368
Comuna 17	104544	9.39	9817
Comuna 18	100339	13.86	13907
Comuna 19	98735	8.75	8639
Comuna 20	65267	12.55	8191
Comuna 21	92336	44.52	41108
Comuna 22	9082	10.87	987
Total	2039626		534974
Total de Afrocolombianos comunas 7, 11, 12, 13, 14, 15, 16 y 21			358402
Porcentaje de Afrocolombianos comunas 7, 11, 12, 13, 14, 15, 16 y 21			66.99

Fuente: Elaboración propia con base en censo básico 2005, DANE.

Anexo B

Comuna	% de población afrocolombiana	Tamaño promedio hogar	% de inasistencia a escolar en diferentes grupos de edad				Nivel educativo (cuatro niveles)			
			3 a 5 años	6 a 10 años	11 a 17 años	18 a 26 años	Ninguna	Especialización Maestría Doctorado	Profesional	Tecnológica
Comuna14	51,2	4,0	49,5	7,1	21,1	87,7	8,3	0,1	1,2	2,0
Comuna15	50,1	4,0	44,8	7,6	19,5	82,2	7,0	0,4	3,1	3,4
Comuna21	44,6	3,9	35,9	4,6	15,6	82,8	4,9	0,3	1,5	2,9
Comuna13	38,1	3,9	42,0	5,6	19,0	83,4	6,4	0,3	2,1	3,1
Comuna 7	37,8	3,6	36,4	4,5	17,6	80,3	5,3	0,3	4,2	4,6
Comuna12	37,5	3,6	35,4	3,3	15,7	80,2	4,7	0,4	4,5	4,5
Comuna11	33,0	3,7	35,6	4,7	15,7	79,0	4,1	0,4	6,3	5,3
Comuna16	27,0	3,6	37,1	4,9	17,2	80,5	4,8	0,5	4,7	3,7
Cali Total	26,4	3,5	35,3	5,2	16,4	74,7	4,4	1,6	9,6	5,0
Comuna 4	22,4	3,4	29,4	5,2	15,9	77,1	3,5	0,6	7,2	5,7
Comuna 5	21,0	3,3	19,5	4,2	11,2	63,8	1,9	1,5	16,2	10,5
Comuna 6	21,0	3,6	30,9	4,0	16,1	76,4	3,6	0,5	5,7	5,7
Comuna 8	19,2	3,4	29,4	4,6	14,9	73,9	3,0	0,7	8,6	6,3
Comuna 9	17,3	3,1	36,1	5,3	16,3	74,0	3,7	0,9	10,4	6,0
Comuna10	15,5	3,5	27,2	4,3	13,0	65,5	2,7	1,7	15,7	7,2
Comuna 3	13,9	2,7	37,9	5,9	19,1	73,1	4,3	2,0	13,2	5,8
Comuna18	12,9	3,5	36,3	5,7	16,5	75,8	4,7	1,0	8,8	4,0
Comuna20	12,6	3,8	45,6	6,6	22,0	86,8	7,4	0,3	2,2	2,2
Comuna 1	12,2	3,6	36,6	4,9	18,5	82,3	5,3	0,5	2,5	2,7
Comuna22	10,7	3,6	15,4	5,5	11,3	57,1	1,6	12,1	27,8	4,3
Comuna 2	10,4	3,1	20,7	4,5	11,4	59,4	1,8	6,3	25,3	7,6
Comuna17	9,5	3,2	11,9	3,5	8,9	46,2	1,2	7,6	31,0	7,0
Comuna19	9,1	3,1	16,1	4,1	11,2	47,4	1,3	6,9	30,6	7,2

Fuente: DANE, Censo 2005. Cálculos: Carlos Viáfara, CIDSE, Universidad del Valle.