

**COMPARABILIDAD DE LOS CENSOS COLOMBIANOS DE 1993 Y 2005:
CAMBIOS EN LA RECOLECCIÓN DE INFORMACIÓN Y DIFICULTADES EN EL
ANÁLISIS DE LAS EVOLUCIONES INTRAURBANAS EN BOGOTÁ.**

Guillaume Le Roux
geógrafo-estadístico
estudiante de doctorado en geografía
UMR Migrinter
guillaumeleroux14@gmail.com
Migrinter, MSHS, bâtiment A5, 5 rue Théodore Lefebvre, 86000 Poitiers, France
(33) 6 50 99 52 58

traducción de francés a español con la colaboración de Antonio Cuello Butelo

RESUMEN

El objetivo de esta ponencia es discutir la comparabilidad de los censos colombianos de 1993 y 2005 en la perspectiva del análisis de evoluciones intraurbanas. Dos cuestiones principales se plantean en este estudio. La primera es la cuestión de la comparabilidad de las definiciones de las unidades censales. A partir del estudio de las definiciones dadas a los encuestadores y del tratamiento de los datos censales de 1993 y 2005 de Bogotá, este trabajo muestra que los cambios más importantes se refieren a la definición de las unidades de vivienda: en 2005, las unidades de vivienda parecen corresponder a entidades más finas que en 1993. La segunda es la validez de resultados obtenidos a partir de la expansión, a un nivel inferior a la localidad (nivel definido por el diseño de muestreo), de características de la población derivados del formulario ampliado de 2005. El estudio de los tamaños de las muestras muestra que sería posible, con factores de extrapolación adecuados, ofrecer resultados al nivel de algunos sectores censales.

INTRODUCCIÓN

El objetivo de los censos nacionales es contabilizar en un momento dado la población y las viviendas en todo el territorio nacional. En la mayoría de los países del mundo se realizan según los mismos principios: " Los cuatro principios fundamentales del censo son la *exhaustividad* (que debe involucrar todos los habitantes de un territorio), el *recuento individual* (que recoge información sobre cada individuo), la *simultaneidad* (la información

recopilada se refiere a la situación en una sola fecha y para hacerlo, la operación debe ejecutarse en unos pocos días) y, finalmente, la *periodicidad* (el censo se debe repetir a intervalos regulares) " (Tabutin, 2006, p.29; traducción personal.).

Otros objetivos pueden ser añadidos, tales como la estimación de indicadores demográficos para la fertilidad, mortalidad y migración que permiten proyectar la evolución de la población, y aportar más datos para responder a preguntas más específicas sobre la actividad y la educación. La producción por división administrativa es también un objetivo fundamental porque permite a los órganos políticos y administrativos mejorar la planificación, y también para establecer los recursos asignados a las autoridades locales (Bussi y Guermond, 2001, p. 106). Los recursos requeridos son considerables, por lo que algunos censos tienen sólo una parte de la información recogida de una manera exhaustiva, otra parte se recoge con una muestra cocensal.

Dentro de las únicas fuentes de datos de población exhaustivos y localizados, los datos censales son un recurso importante para la investigación. La producción de estos datos a intervalos regulares de tiempo permite estudios de especial interés. Sin embargo, este tipo de estudio es sensible a los cambios en la ejecución de los censos: cambios de definiciones de las unidades estadísticas (Toulemon, 2011), cambios en la manera de recopilar la información (Courgeau, 1999), calidad de la cobertura geográfica de la recolección exhaustiva. A un cierto nivel de agregación, a través del estudio de las dinámicas demográficas o del uso de fuentes secundarias, ciertos ajustes permiten controlar o armonizar los datos entre dos censos (Flórez, 2000). Pero ¿qué pasa a una escala intraurbana?

El objetivo de esta ponencia es discutir la comparabilidad de los censos colombianos de 1993 y 2005 en la perspectiva del análisis de evoluciones intraurbanas en Bogotá. Dos cuestiones principales se plantean en este estudio: i) ¿Hasta que punto son comparables las unidades

censales (viviendas, hogares, personas) definidas por los censos de 1993 y 2005? ii) ¿En que medida el formulario ampliado del censo de 2005 (aplicado a una muestra de hogares) permite un estudio intraurbano preciso? ¿Es razonable o no producir resultados a un nivel de agregación inferior al definido por el diseño de la muestra, es decir, las localidades de Bogotá?

Esta ponencia se divide en tres partes. La primera presenta de manera general los censos colombianos de 1993 y 2005 mediante la exposición de los elementos que pueden dificultar un análisis de las evoluciones intercensales o un análisis a escala intraurbana. La segunda parte examina la cuestión de la comparabilidad de los censos a través del estudio de las definiciones y de los recuentos de las unidades censales. La tercera responde a la cuestión de la validez de un análisis intraurbano fino a partir de datos de la muestra cocensal en 2005¹.

1. LA CUESTIÓN DE LA COMPARABILIDAD ENTRE LOS CENSOS DE 1993 Y 2005

La cuestión de comparabilidad es más sensible cuando se aplica a unos datos intraurbanos por dos razones, en primer lugar las encuestas post-censales para evaluar la calidad de la recolección no pueden recuperar los datos a nivel fino, y segundo un cambio en la manera de recoger la información tendrá importantes repercusiones.

1.1 Diferencias en la recolección de la información

Los objetivos declarados y la manera de recopilar los datos son muy importantes ya que determinan el tipo y la calidad de la información recopilada. Los objetivos en términos de

¹ La muestra cocensal es una muestra de hogares seleccionada en tiempo real durante el operativo del Censo 2005

producción de información son en gran parte comunes entre 1993 y 2005: contar la población del país, recolectando sus características demográficas y socioeconómicas, recoger las características de viviendas y hogares, actualizar la cartografía urbana y rural, y producir información espacializada en el contexto de la descentralización a nivel nacional, regional, departamental, municipal (diferenciando entre áreas urbanas y rurales) (DANE, 1996, p.10; DANE, 2009, p.23-24). En 2005, nuevos objetivos se añaden para capturar la ubicación y características de los establecimientos económicos y las unidades agropecuarias y para mejorar la estratificación, con fines de cobros de servicios públicos domiciliarios.

En términos de población y viviendas, los objetivos son similares, sin embargo, a la hora de producir información para cada división territorial, su aplicación es distinta dado que la recolección de datos es totalmente exhaustiva en 1993, mientras que en 2005 una parte es recogida por una encuesta aplicada a una muestra de hogares, con la intención de dar representatividad a nivel de las localidades en las grandes ciudades. Esta opción, obviamente, tiene consecuencias en el análisis intraurbano porque la información recogida por esta encuesta se limita a la división en localidades de las principales ciudades.

Otras diferencias en la forma de recolección pueden crear desigualdades de calidad en la información producida. Desde 1985, el censo es un censo "de jure", es decir que la población se cuenta en su lugar de residencia habitual, pero la duración de recolección es distinta. En 1993, en las zonas urbanas, el censo se llevó a cabo en el día el 24 de octubre (DANE, 1996, p.12). Mientras que en 2005, la recolección duró un año, desde el 22 de mayo 2005 hasta el 22 de mayo de 2006, dividida en diferentes fases. Para garantizar un mejor control y reducir al mínimo el sesgo de recolección, cada fase implica un conjunto de municipios clasificados en grupos de dinámica poblacional y de migraciones similares a partir de datos de los censos de

1985 y de 1993, y de otros datos auxiliares. Todo el área metropolitana de Bogotá fue censado en la misma fase (DANE, 2009, p.37-41), que duró un poco más de un mes (DANE, 2008, p.19). El DANE presenta este cambio en la duración de recolección como un avance, porque permite un mejor control de riesgos, disminuye los costos y mejora la cobertura y la calidad de información, recogida por los investigadores a través de un mayor tiempo de aprendizaje (DANE, 2007-a, p.32-36). Estos argumentos pueden ser debatidos, como lo fueron cuando se estableció el censo continuo en Francia². Un censo continuo plantea muchos problemas técnicos y teóricos (Courgeau, 1999). Los cambios de residencia, los nacimientos y los fallecimientos que se producen durante todo el año son difíciles de tomar en cuenta. El riesgo de doble cuenta potencial aumenta, así como las posibles omisiones. Individuos altamente móviles se pueden contar varias veces al año. Las migraciones son particularmente importantes y afectan a ciertos grupos de edad. "Su distribución [no es] consistente en el tiempo [o] en el espacio" (Courgeau, 1999; traducción personal). El estudio de la dinámica poblacional de los municipios, sin embargo, permite ajustar un poco el censo de población, pero este ajuste, a base de datos auxiliares no resuelve todas las inexactitudes introducidas. Existen preguntas de control para saber si la persona ya fue censada en otro lugar. Sin embargo, "queda la posibilidad que personas desplazadas que cuenten con parte del hogar en el lugar de origen hubiesen sido referidas por sus familiares y éstos a su vez sin conocimiento se hayan empadronado nuevamente" (DANE, 2007-b, p.22) o "que un volumen de población no sea captada en el censo, dado que al realizarse en un municipio estas personas se muevan hacia municipios ya censados" (DANE, 2009, p.40). Además, la definición del momento censal (día de referencia para definir las variables) es el día de inicio del operativo de

² Ver el debate « Recensement de la population en continu : choisir le temps et l'espace ? », *Cybergeo*, <http://cybergeo.revues.org/66>

recolección para cada fase, es decir que cada grupo de municipios tiene un momento censal diferente.

El debate no es sencillo y respecto a la cuestión intraurbana, la duración de recolección de 5 semanas en 2005 para la fase de las grandes ciudades, en cierta medida limita estos sesgos.

1.2 Evaluación de los censos : la cobertura

Después de la recolección, los datos censales son sometidos a una evaluación de la población omitida para ajustar los resultados: directamente a través de una encuesta de cobertura, o de forma indirecta mediante el análisis de la dinámica demográfica (DANE, 1981; Manrique De Llinás, 1990; DANE, 1996; DANE, 2007-c). Las tasas de cobertura estimadas (TABLA 1) son alrededor del 90% de los censos de 1973, 1985 y 1993, y significativamente más altas para el año 2005. Según el DANE, la actualización de la cartografía, la introducción de nuevas tecnologías para garantizar la recogida de información, el desarrollo de nuevos métodos de los programas de gestión de personal, la aplicación de ediciones automatizadas y la imputación de datos faltantes serían las razones de una recolección de más alta calidad en 2005 (DANE, 2007-c, p.2). Sin embargo, esta afirmación y las tasas de cobertura estimadas, extraordinariamente altas, pueden ser impugnadas a varios puntos. En primer lugar, la duración de la recogida, de hecho, aumenta el riesgo de doble cuenta y omisiones. Además, los procedimientos de corrección, edición e imputación de datos (DANE, 2007-a, p.236-239 y p.319-335) parecen sobreestimar la cobertura y la calidad de la información. La regularidad excepcional de la pirámide de edades y las tasas de masculinidad sugieren que se hicieron correcciones a los datos demográficos (Dureau y Flórez, 2008, p.4). Estos cambios son el resultado de las imputaciones de los valores omitidos, de correcciones de inconsistencia (DANE, 2008) y de los procesos de conciliación censal (DANE, 2007-c, p.3-14). Por otra parte, la tasa de cobertura estimada en 2005 no se deriva directamente de una encuesta de

cobertura, sino de datos específicos del SISBEN (Sistema de Identificación de Potenciales Beneficiarios de Programas Sociales), lo que puede ser objeto de controversia (Mallarino, 2007). Esta aparente mejor calidad oculta la manipulación sustancial de los datos censales y no tiene en cuenta las dificultades encontradas por los encuestadores.

TABLA 1: tasas de cobertura de los censos de 1973, 1985, 1993 et 2005

Fechas	Tasa de cobertura		Fuentes
	Bogota	Cundinamarca	
1973	90,70%	94,70%	DANE (1981) <i>XIV censo nacional de población y III de vivienda, Octubre de 1973</i> , Bogota, DANE, 490 p.
1985	92,40%	89,60%	DE LLINAS MANRIQUE H. (1990), <i>La población de Colombia en 1985: Estudios de evaluación de la calidad y cobertura del XV Censo Nacional de Población y IV de Vivienda</i> . Bogotá, DANE, 178 p.
1993	90,20%	88,40%	DANE (1996) <i>Censo 1993. Resumen nacional</i> , Bogota, DANE, 163 p.
2005	98,50%	97,20%	DANE (2007) <i>Conciliación censal 1985-2005. Colombia. Proceso de conciliación censal 1985-2005</i> . Bogotá, DANE, 235 p.

Para concluir, la especificidad la más importante del censo de 2005, que puede interferir con los estudios intraurbanos finos en ciertas características de la población y de las viviendas, es el hecho que una parte de la información se recoge de manera exhaustiva, mientras que otra parte se capta a partir de una muestra.

1.3 La información recopilada : el censo básico y la muestra cocensal

Los distintos censos no recogen ni la misma cantidad de datos, ni el mismo nivel de información; teniendo incluso información para la cual las definiciones varían. Sin embargo, podemos observar una similitud de la información y de los temas.

En 1993 como en 2005, los datos recogidos de manera exhaustiva son:

- *vivienda*: tipo de vivienda, conexión a los servicios, número de hogares
- *hogar*: número de personas, equipamiento, tipo de financiación de la vivienda, tenencia del hogar

- *miembros del hogar*: características sociodemográficas (sexo, edad, estado civil, migración con una recolección ampliada en 2005), educativas, económicas (ocupación) y de fecundidad

El formulario ampliado de 2005, aplicado en la encuesta cocensal, captura información adicional que puede ser interesante, como los ingresos de los hogares, los medios de transporte de los miembros del hogar, sus lugares de trabajo y de estudio, e informaciones de salud. Sin embargo, a diferencia del censo de 1993, que recoge todo de manera exhaustiva, algunas de las características claves para estudiar las condiciones de alojamiento o la situación socioeconómica de los miembros del hogar se introducen sólo a través de la encuesta cocensal, como la tenencia del hogar, el número de cuartos y dormitorios del hogar, informaciones sobre la actividad y el tiempo de trabajo de los miembros del hogar.

La comparabilidad de las variables de ciertos temas depende de la precisión en la recogida (a través de categorías) y de las definiciones dadas. Sin entrar en el debate sobre las categorías utilizadas en los dos censos y las cuestiones de comparabilidad que surgen, nos centramos aquí en las preguntas básicas acerca de la comparabilidad de las unidades censales, a través de sus definiciones en ambas fechas y de la aplicación efectiva de estas definiciones.

2. LA COMPARABILIDAD DE LAS UNIDADES CENSALES

La cuestión de la comparabilidad de las definiciones de las unidades censales se trata tanto teóricamente, a partir de las instrucciones dadas a los encuestadores, como empíricamente, a partir del tratamiento de los datos censales de 1993 y 2005 en la zona urbana del Distrito de Bogotá, estudiando los recuentos de las diferentes unidades censales (personas, hogares, vivienda) y sus combinaciones (número de individuos en el hogar, número de hogares en la vivienda).

2.1 La comparación de las definiciones

Durante cada censo, los formularios son diferentes para los hogares colectivos, los hogares particulares y las zonas indígenas. Aquí nos centramos sólo en el formulario de los hogares particulares que es el formulario el más común.

El conteo de las unidades estadísticas (vivienda, hogar, persona) depende de la definición en cada censo, de las instrucciones dadas a los encuestadores y de la calidad de su capacitación. La definición de la vivienda es la misma en ambas fechas (TABLA 2). Las diferencias radican tanto en la definición de los hogares, como en la definición del miembro del hogar, definido en relación con su residencia habitual. La principal diferencia en la definición del hogar es el fortalecimiento de ésta como una unidad económica. Los censos no están de acuerdo en la definición del miembro del hogar: en 1993 los miembros del hogar son los que están presentes en el momento de la investigación o ausentes, pero sin otra residencia en otro lugar; mientras que en 2005, los miembros del hogar son los que viven la mayor parte del tiempo en la vivienda.

TABLA 2: definiciones de las unidades de observación para 1993 y los censos de 2005

Fecha	Vivienda	Hogar	Residente habitual
1993	“espacio independiente o separado con áreas de uso exclusivo, habitado o destinado por una o más personas”	“persona o grupo de personas, parientes o no, que viven bajo un mismo techo y generalmente comparten los alimentos”	“[personas] quienes hace parten del hogar y se encuentran presentes en el momento del censo o quienes viven en el hogar, pero en el momento del censo se encuentran ausentes dentro o fuera del país y que no tienen residencia en otra parte”
2005	“espacio independiente o separado con áreas de uso exclusivo, habitado o destinado por una o más personas”	“persona o grupo de personas, parientes o no, que ocupan la totalidad o parte de una vivienda; atienden necesidades básicas con cargo a un presupuesto común y generalmente comparten las comidas”	“persona que vive permanentemente o la mayor parte del tiempo en una vivienda, aunque en el momento de la encuesta se encuentre ausente”

Fuentes: DANE (1996, p.161), DANE (2007, p.424-444)

Podemos observar que, a pesar de las diferencias de definición, su aplicación sobre el terreno puede dar datos muy próximos (hogar, residente habitual), y que, aún siendo la misma definición en ambos censos, los datos recogidos para la vivienda pueden diferir de manera ostensible. A través de las instrucciones dadas a los encuestadores, podemos ir un paso más allá, mediante el examen de casos particulares y cómo los diversos censos les consideran.

La vivienda

Las definiciones son las mismas en 1993 y 2005 y se basan en un acceso independiente. Las principales dificultades radican en la definición de unidades de vivienda en el caso de viviendas subdivididas, incluyendo el caso especial de los inquilinatos³.

En 1993, el manual del encuestador no parece presentar los problemas que podrían encontrar los encuestadores para determinar las unidades de vivienda (DANE, 1993). La aplicación de la definición estricta conduce a considerar las diferentes unidades independientes de los inquilinatos como viviendas de tipo “cuarto” desde el momento en que se puede entrar en la vivienda por una zona de circulación común, como el patio, o que tiene acceso directo a la vía pública. En 2005, el caso especial de los inquilinatos se especifica explícitamente como un conjunto de viviendas de tipo “cuarto” (DANE, 2005, p.56), a condición de que cada unidad tenga una entrada independiente en la vía pública o en un espacio de circulación común.

La forma de contar las viviendas parece ser constante a través de los censos de 1993 y 2005, con poca ambigüedad en el conteo de las unidades en inquilinatos o en viviendas subdivididas.

³ « Un **inquilinato** es una edificación adaptada o transformada para alojar varios hogares que comparten servicios y con una entrada común desde la calle. Cada hogar vive generalmente en uno o más cuartos. Los servicios no son exclusivos y en la mayoría de los casos están ubicados en el patio, solar o corredores » (DANE, 2007-d)

El hogar

Además de compartir las comidas, el censo de 2005 define el hogar de acuerdo a un presupuesto común para apoyar las necesidades básicas. Sin embargo y como en 1993, la pregunta en el formulario se refiere a la enumeración de los grupos de residentes habituales que cocinan por separado.

El manual de los encuestadores dice que los individuos que pagan por el derecho a estar en la vivienda y por los alimentos, se consideran parte del hogar, como empleados domésticos que viven la mayor parte del tiempo en la vivienda (DANE, 2005, p.53). Las personas que alquilan una o dos habitaciones en una vivienda y preparan comida o comen por separado, son hogares distintos. En los casos donde grupos de personas no comparten los alimentos, pero comparten los gastos de arriendo y servicios, se consideran parte de un mismo hogar, como estudiantes compartiendo una vivienda en arriendo. Este último caso se diferencia del censo de 1993 en el que podría considerarse como hogares separados.

Los residentes habituales del hogar

La inclusión de ciertos tipos de residentes habituales en el hogar puede causar problemas para contar el número de hogares, pero especialmente en el estudio comparativo del tamaño y de la composición de los hogares. En 1993 y 2005, no se cuentan como miembros del hogar las personas que están pagando condenas en cárceles, las que residen en cuarteles del Ejército, de la Fuerza Aérea o de la Marina, las personas en internados de estudio, asilos, conventos, monasterios o campos de trabajo (DANE, 1996, p.152; DANE, 2005, p.54). La población está inmovilizada en 1993 y 2005 en su lugar de residencia. Los presentes que viven de forma permanente en la vivienda y los empleados domésticos y sus familias son contados. Las diferencias radican en las ausencias en el momento censal y los casos de multi-residencia. En 1993 y 2005, son contadas las personas temporalmente ausentes que no tengan residencia en

otro lugar, como los agentes viajeros, personas en viaje de vacaciones o de negocios (si la duración es de 6 meses o menos para el censo de 2005 y sin duración definida en 1993), los individuos internados en hospitales o clínicas y los marinos mercantes. Varios casos especiales sólo aparecen en el manual de 2005: las personas secuestradas, las personas desplazadas (independientemente del tiempo de permanencia en la vivienda), las personas detenidas en inspecciones de policía, las personas que prestan el servicio militar y que duermen en la vivienda son contadas.

Los casos de multi-residencia están claramente definidos en 2005, pero no en 1993: parece que un multi-residente se cuenta en 1993 donde se encuentra el día del censo, mientras que en 2005 está censado en el lugar donde pasa la mayor parte del tiempo.

2.2 Ilustración de los problemas de comparabilidad entre 1993 y 2005

Los recuentos de las unidades censales en 1993 y 2005 en el área urbana de Bogotá y la combinación de estas unidades (número de hogares que comparten sus viviendas y tamaño de los hogares) plantean algunas preguntas (TABLA 3).

El crecimiento de los hogares es más rápido que el crecimiento de la población y más lento que el crecimiento de las viviendas. Dos tendencias son observables. La primera, que explica el rápido crecimiento de los hogares con respecto a la población, es la reducción de los tamaños de los hogares, principalmente debido al aumento de la proporción de hogares unipersonales y una reducción relativa de los hogares de tipo compuesto (familia nuclear con otros parientes) o extenso (familia nuclear y personas con ninguna relación de parentesco). La segunda pone de manifiesto la disminución del número de hogares por vivienda. En 1993, casi la mitad de los hogares compartieron sus viviendas. La reducción de esta práctica es muy importante durante el periodo de estudio y en 2005 sólo el 20% de los hogares compartían sus viviendas. ¿ Esta disminución es sólo debido a una atenuación de la crisis de la vivienda?

Las diferencias entre 1993 y 2005 son tan fuertes que uno puede sospechar una diferencia en el recuento de las unidades de hogar y / o vivienda. El cambio en la definición del hogar (personas que comparten los gastos de arriendo y servicios, incluso si comen por separado), puede sugerir que esto se debe a un recuento diferente de los hogares. Estos casos especiales no parecen explicar por sí solos este fenómeno, al contrario el crecimiento de hogares unipersonales y la reducción de hogares de gran tamaño muestran la dinámica inversa. La comparación de las cifras de crecimiento de viviendas en el período 1993-2005 con un estudio de las formas de producción de la vivienda en Bogotá (Cuervo, 2008) tiende a la discusión de una diferencia en el recuento de las unidades de vivienda. Hay un crecimiento de 817,803 viviendas entre 1993 y 2005 según los censos. El estudio de las licencias de construcción muestra que la producción de vivienda formal en el mismo período es de 376,862 unidades (Cuervo, 2008, p.38) o el 46% del crecimiento de viviendas censadas. Un 54% de las viviendas serían producidas de manera informal o por subdivisión de las viviendas existentes, una proporción que supera todas las estimaciones de la autoconstrucción desde 1928 en los períodos intercensales. La explicación se podría buscar en los cambios de enumeración de las unidades de vivienda.

TABLA 3: evolución de las recuentos de las unidades censales entre 1993 y 2005 en Bogotá (zona urbana)

Unidad	Indicador	1993	2005
Viviendas	Número de viviendas	940 489	1 758 292
	Tasa media de crecimiento anual	5,4 %	
Hogares	Número de hogares	1 251 446	1 927 390
	Tasa media de crecimiento anual	3,7 %	
Población	Número de personas	4 906 585	6 725 493
	Tasa media de crecimiento anual	2,7 %	
Tamaño de los hogares	% hogares unipersonales	8,4 %	13,3 %
	% hogares de 6 personas y mas	17,2 %	9,9 %
	Número promedio de individuos por hogar	3,9	3,5
Hogares que comparten sus viviendas	% hogares que comparten sus viviendas	47,4 %	20,2 %

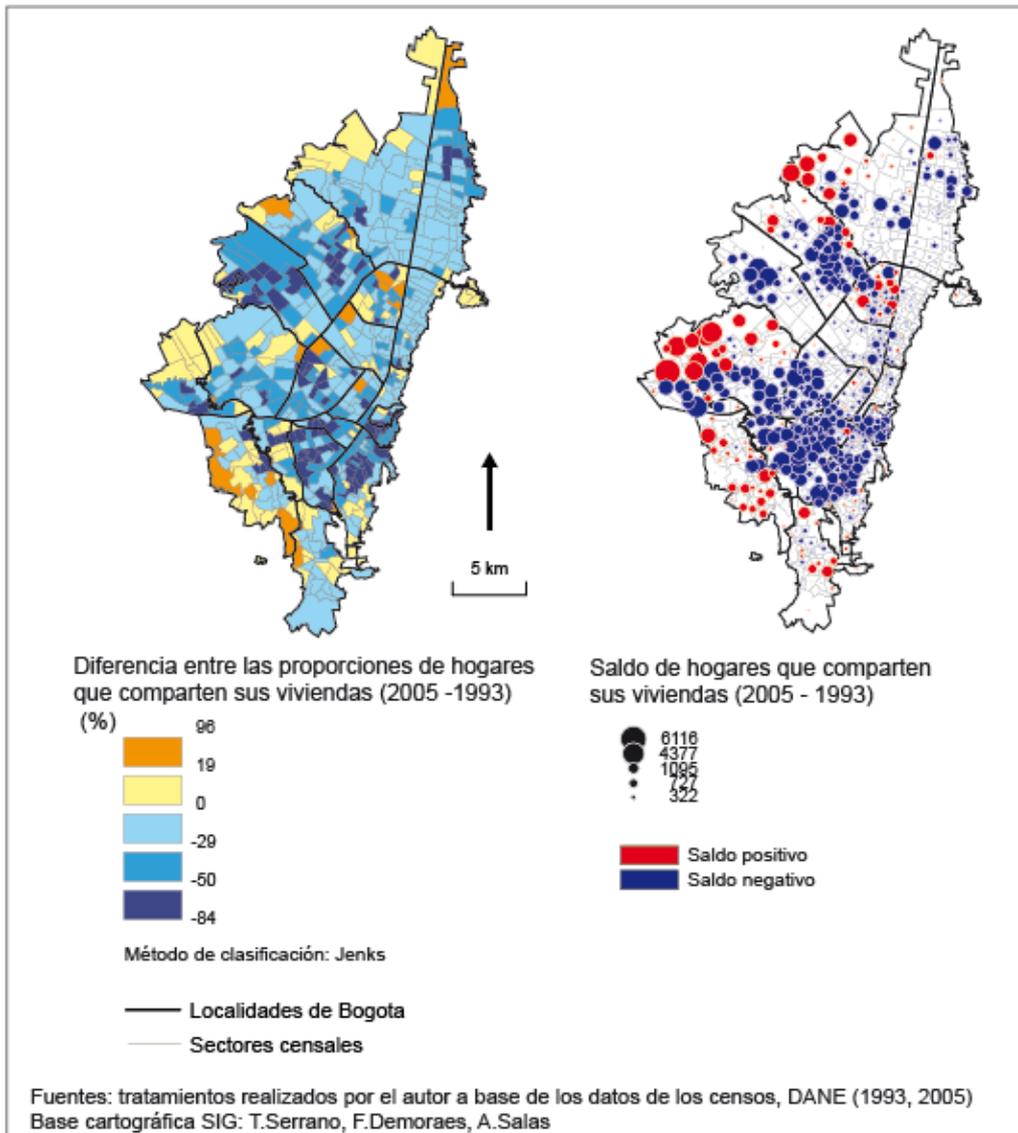
Fuentes: procesamientos realizados por el autor a partir de los censos, DANE, 1993 y 2005

¿Una disminución real de la vivienda compartida?

Las fuertes diferencias en la práctica de compartir vivienda entre 1993 y 2005 se observan en el pericentro sur y en una parte de la periferia oeste más cercana al centro, con una disminución a veces mayor de 50 puntos en la proporción de hogares que comparten sus viviendas en 1993 y 2005 (FIGURA 1). Estas dos zonas se caracterizan también por un crecimiento bajo y a veces negativo de la población (FIGURA 2), y un incremento relativamente bajo del número de hogares, que puede ser explicado en parte por el aumento de los hogares unipersonales. La población en estas zonas es relativamente estable y esas diferencias parecen sorprendentes.

FIGURA 1

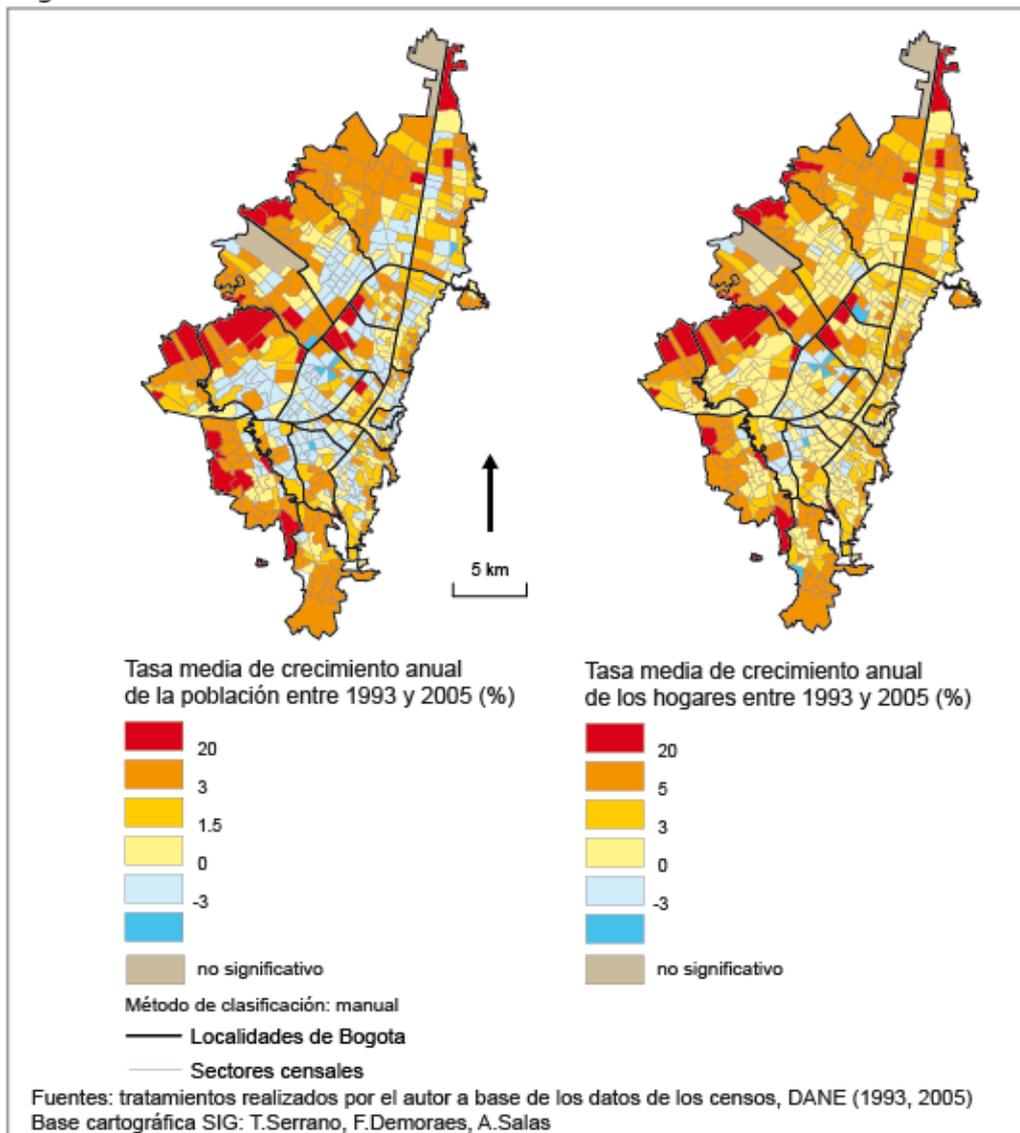
Evolución de los hogares que comparten sus viviendas. Bogota 1993 y 2005



G.Le Roux, 2011

FIGURA 2

Tasa media de crecimiento anual de la población y de los hogares entre 1993 y 2005
Bogota



G.Le Roux, 2011

Desaparecen numerosas viviendas compartidas por muchos hogares

Esta disminución de las viviendas compartidas se ve no sólo en los porcentajes, sino también en gran medida en el número de viviendas. Hay una ruptura en la práctica de compartir la vivienda entre dos, y tres o más hogares, las viviendas ocupadas por tres hogares o más tienden a desaparecer (TABLA 4).

TABLA 4: número de hogares por vivienda en Bogotá (1993 y 2005)

		1 hogar	2 hogares	3 hogares	4 hogares	5 hogares	6 y mas	Total
1993	Número de viviendas	669 204	105 007	49 219	23 606	11 214	10 474	868 724
	Porcentaje	77,0%	12,1%	5,7%	2,7%	1,3%	1,2%	100,0%
2005	Número de viviendas	1 538 875	107 948	28 497	10 722	4 531	3 133	1 693 706
	Porcentaje	90,9%	6,4%	1,7%	0,6%	0,3%	0,2%	100,0%

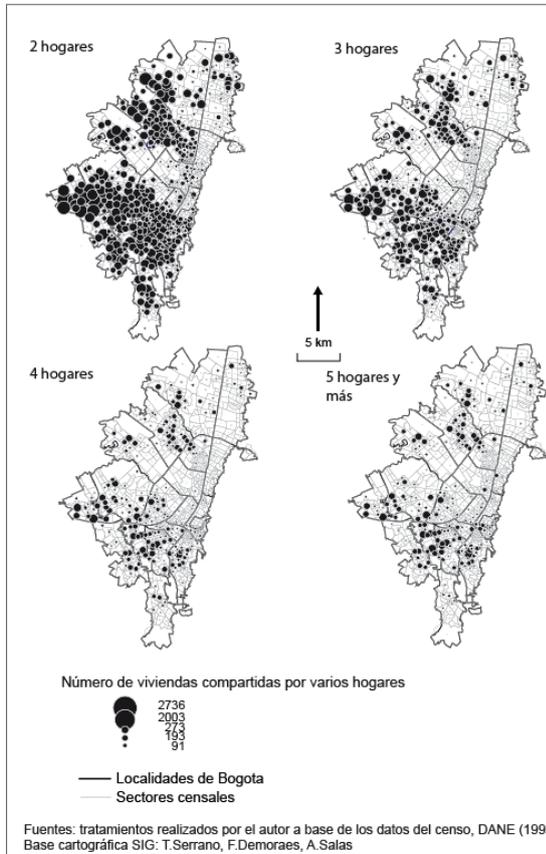
Fuentes : DANE, 1993 et 2005

La disminución entre 1993 y 2005 no es uniforme en el espacio (FIGURA 3). Las periferias extremas del oeste y del sur, así como la localidad de Barrios Unidos, están experimentando un aumento en las viviendas compartidas por dos hogares, como consecuencia del crecimiento de la población en estas zonas y de la continuidad de esta práctica. Al mismo tiempo se observa una ruptura con las viviendas compartidas por más de tres hogares, las cuales disminuyeron drásticamente, salvo en unos pocos lugares de la periferia como Bosa o en las periferias lejanas de Suba.

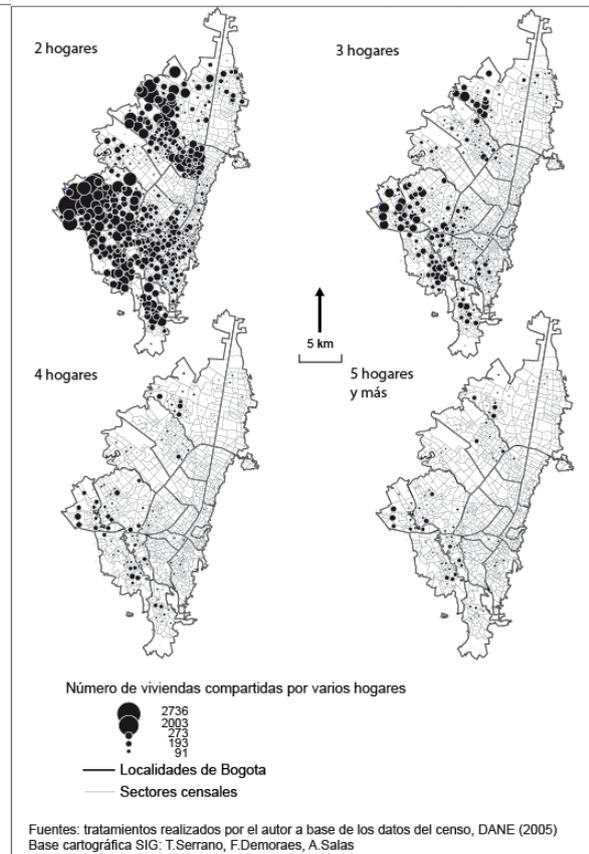
Con respecto a las cercanías de la periferia oeste y el pericentro sur, la disminución de la práctica es neta para las viviendas compartidas por dos hogares y hay pocas viviendas compartidas por más de dos hogares en 2005. Esta fragmentación de la evolución, con la desaparición de muchas viviendas compartidas por más de dos hogares, parece antinatural. ¿Estas formas de convivencia desaparecen realmente o es el resultado de una diferencia en el recuento de las unidades censales? La respuesta es, probablemente, en ambas cosas, pero la desaparición de la práctica por sí sola parece poco probable.

FIGURA 3

Número de viviendas compartidas por varios hogares. Bogota 1993



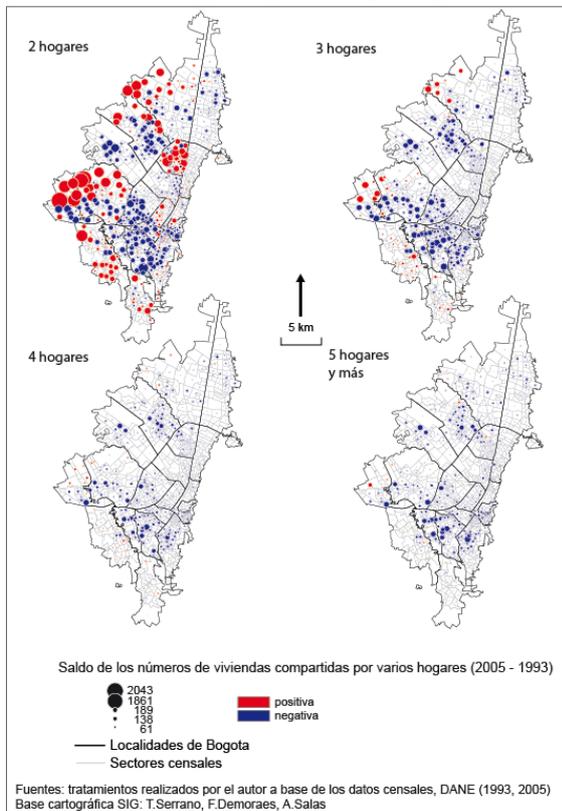
Número de viviendas compartidas por varios hogares. Bogota 2005



G.Le Roux, 2011

G.Le Roux, 2011

Saldo de las viviendas compartidas por varios hogares. Bogota 1993-2005



G.Le Roux, 2011

¿Una disminución en la práctica de compartir su vivienda puede ser explicada por cambios en la composición de la población?

La mejora en las condiciones de vida, los cambios en la composición de la población y la transformación de las viviendas podrían explicar una disminución real en la práctica de compartir su vivienda. Para explorar esta pista, se construye un modelo para explicar la proporción de hogares que comparten sus viviendas en 1993 a partir de variables que describen la población y las viviendas. Las estimaciones realizadas por este modelo se utilizan para "predecir" las proporciones de hogares compartiendo sus viviendas en 2005, teniendo en cuenta los cambios en la composición de la población y en las características de las viviendas. La hipótesis que fundamenta este enfoque es que las variables asociadas a la práctica de compartir su vivienda no cambian entre 1993 y 2005, es decir, que la composición de la población, las características de los hogares y de las viviendas determinan de la misma manera la proporción de hogares que comparten sus viviendas en 1993 y 2005. Esto supone que esta práctica es típica de un cierto tipo de población y de una cierta configuración de las viviendas.

Puesto que los datos desagregados no están disponibles para 2005, hemos construido este modelo a nivel de los sectores censales. La idea es modelizar la proporción de hogares que comparten sus viviendas en 1993, agrupada en grandes categorías, con una regresión logística ordinal para "predecir" las clases correspondientes en 2005. La comparación de las clases previstas y observadas por el censo puede mostrar inconsistencias. Se supone que la práctica de compartir la vivienda se relaciona con los tipos de viviendas encuestadas, el nivel socioeconómico, el tipo de tenencia de los hogares, la composición del hogar y la edad de los individuos. Las variables incluidas en el modelo después de un proceso de selección de

variables son: la proporción de casas, apartamentos y cuartos; la proporción de jefes de hogar de nivel educativo secundario y superior; la proporción de hogares unipersonales y nucleares; la proporción de individuos menores de 15 años, entre 20 y 34 años y de 60 años o más; la proporción de hogares en arriendo y propietarios.

La capacidad de predecir del modelo resultante puede estimarse a partir de la matriz de confusión, que compara las clases estimadas con las clases observadas (TABLA 5). Para cada clase, más de la mitad de las observaciones son predichas correctamente por el modelo. La clase "más de 55%" se predice con el 85% de las observaciones bien clasificadas, y no hay confusión con la clase "Menos de 15%." Los errores observados son en gran parte cambios entre las clases contiguas.

TABLA 5: matriz de confusión del modelo de vivienda compartida en 1993

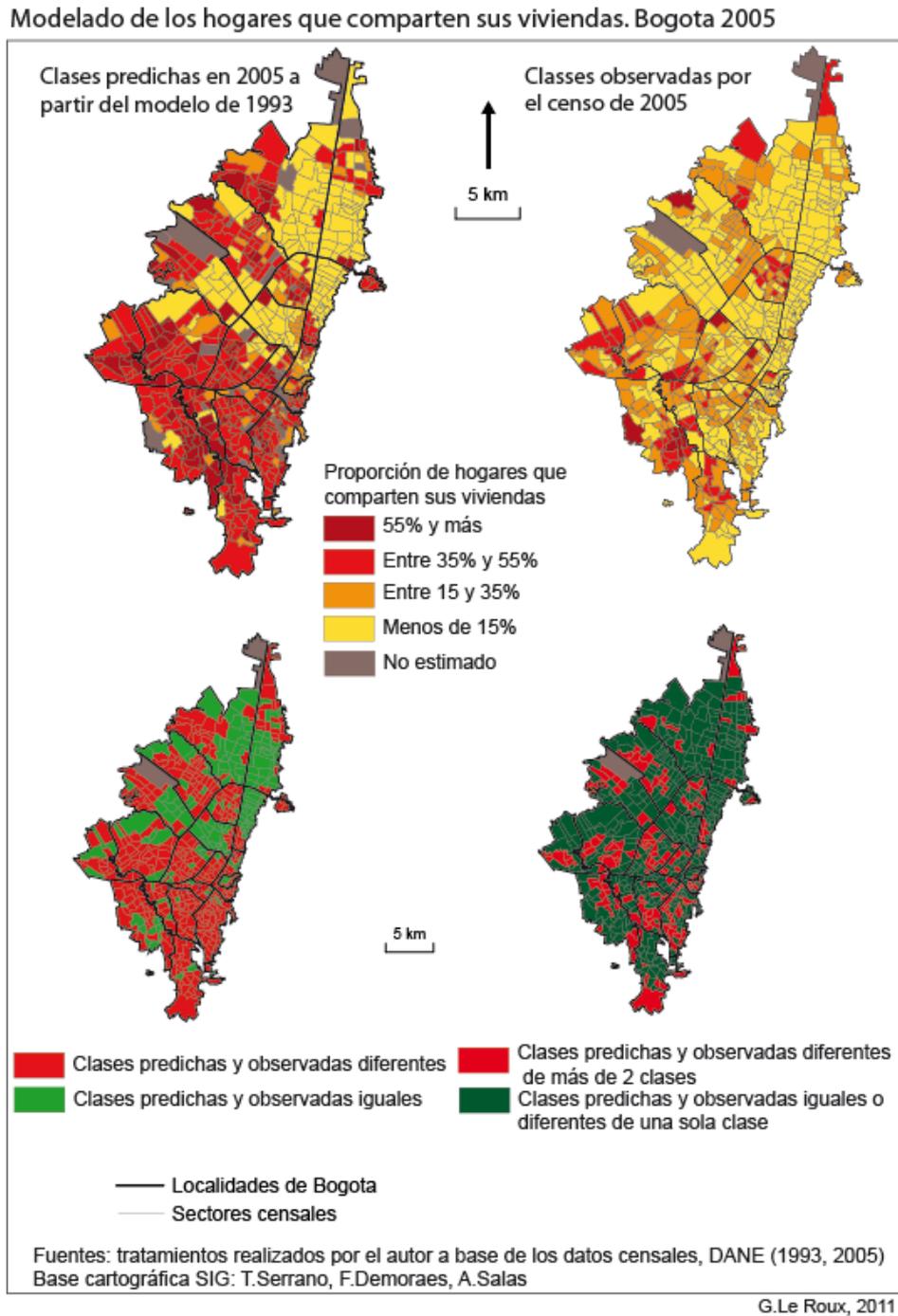
Clases estimadas		Clases observadas			
		Menos de 15%	Entre 15 y 35%	Entre 35 y 55%	Más de 55%
Menos de 15%	<i>N</i>	93	23	4	0
	<i>% col</i>	64,14	19,83	3,01	0
Entre 15 y 35%	<i>N</i>	34	64	15	5
	<i>% col</i>	23,45	55,17	11,28	2,14
Entre 35 y 55%	<i>N</i>	6	22	93	12
	<i>% col</i>	4,14	18,97	69,92	5,13
Mas de 55%	<i>N</i>	12	7	21	199
	<i>% col</i>	8,28	6,03	15,79	85,04
Ambigüedad de la clasificación	<i>N</i>	0	0	0	18
	<i>% col</i>	0	0	0	7,69
Total	<i>N</i>	145	116	133	234
	<i>% col</i>	100	100	100	100

Fuentes : procesamientos realizados a partir del censo de 1993 del DANE

La aplicación de este modelo a los datos de 2005 permite destacar varios fenómenos (FIGURA 4). En primer lugar, las tendencias generales muestran que las proporciones observadas son más bajas que las proporciones estimadas, pero reflejan las mismas tendencias. Este resultado es importante porque muestra que, aun si las proporciones predichas no son las mismas que las proporciones observadas, las características de la

población, de los hogares y de las viviendas utilizadas para construir el modelo son relacionadas con la proporción de hogares que comparten sus viviendas en 2005. Casi todos los sectores estimados en la clase "Menos de 15%" corresponden a las observaciones reales, lo que refleja la tendencia del modelo a estar por encima de los valores observados. Las diferencias se observan principalmente en la mitad sur y en la periferia oeste del área urbana de Bogotá. En estas áreas, hay que señalar que cuando las proporciones de hogares que comparten sus viviendas son altas, las predicciones son buenas, como en Bosa y en Ciudad Bolívar. Errores de más de 2 clases (que corresponden a diferencias de por lo menos un 20% entre la predicción y lo observado; con valores observados casi siempre inferiores a la predicción) se encuentran principalmente en el pericentro y las periferias cercanas sur y oeste. En su mayoría, corresponden a áreas donde la proporción observada es inferior al 15% y a los sectores que han experimentado el más fuerte descenso en la práctica de compartir la vivienda entre 1993 y 2005. Por lo tanto, estas disminuciones no se pueden explicar sólo por una transformación de la población y de las viviendas en estas áreas.

FIGURA 4



Este análisis muestra incoherencias que aparecen localizadas. Sería interesante ver si esta dinámica espacial está relacionada con las áreas asignadas a los encuestadores en cargo de la enumeración. Además de una diferencia de enumeración entre el censo de 1993 y el censo de

2005, se pueden añadir diferencias de enumeración entre encuestadores de un mismo año. Sin embargo, no tenemos los datos necesarios para responder esta pregunta.

En conclusión, algunos fenómenos, que no son fácilmente mesurables, pueden explicar en cierta medida la disminución de las viviendas compartidas: la evolución de las relaciones entre los individuos y su hábitat, el aumento del nivel de vida en general, la subdivisión de las viviendas existentes, el aumento de la oferta de vivienda. Sin embargo los análisis anteriores muestran que hay una diferencia en la manera de contar las unidades de vivienda entre 1993 y 2005.

3. EL ANÁLISIS INTRAURBANO Y LA MUESTRA COCENSAL EN 2005: ¿QUE PRECISIÓN GEOGRÁFICA?

En 1993, aún cuando el objetivo es generar información a nivel municipal (urbano y rural), la exhaustividad de la recolección permite producir resultados a todas las divisiones geográficas, por ejemplo a nivel de las manzanas de Bogotá. En 2005, algunas informaciones se han recogido de forma exhaustiva y otras sólo a través de una encuesta aplicada a una muestra de hogares seleccionados durante el operativo de recolección exhaustiva. Esto puede ser un obstáculo para un análisis infra-metropolitano. Las informaciones recogidas por la encuesta limitan la producción de resultados a un cierto nivel de división administrativa: la validez de la extrapolación de los datos en un determinado nivel de división administrativa depende del diseño de muestreo aplicado. El formulario ampliado de 2005 se aplicó a una muestra de hogares con diseño de muestreo definido para asegurar una representatividad a nivel de las localidades de Bogotá, y no de los sectores censales (DANE, 2009). Sin embargo, las variables de localización geográfica permiten producir información en una escala más fina. El estudio de las tasas de muestreo y de los tamaños de las muestras ayuda a dar elementos de

respuestas sobre la representatividad y la precisión de las estimaciones de las características de la población a un nivel inferior (sectores censales).

3.1 ¿Un muestreo adaptado sólo a nivel de las localidades?

En 2005, una parte de la información fue recogida de manera exhaustiva (formulario básico) y otra parte a través de una encuesta aplicada a una muestra de hogares seleccionados al azar (formulario ampliado). Como vimos en el primer capítulo, algunas características captadas por la encuesta pueden ser necesarias para ciertos estudios intraurbanos. El formulario ampliado fue aplicado a cerca del 12% de los hogares del territorio nacional, que fueron elegidos según una selección de Bernoulli, con una estratificación urbana/rural por municipios, y por localidades en las grandes ciudades. Las tasas de muestreo varían de 1,2% a 60%, dependiendo de la población del municipio y de su distribución entre las zonas urbanas y rurales (Dureau y Flórez, 2008). El tamaño de la muestra por municipio (o localidad para las grandes ciudades) se determinó con el fin de garantizar precisiones de las estimaciones equivalentes para cada uno de ellos: el número mínimo de hogares encuestados se calcula para garantizar la estimación de una característica presente en un 10% de la población, con un error de muestreo⁴ de 7% o menos.

Las tasas de muestreo (el número de hogares encuestados dividido por el número de hogares censados) por localidad en el Distrito de Bogotá son muy variables de una localidad a otra, entre el 1,8% y el 40,7% (TABLA 6). Las tasas de muestreo son más bajas en las localidades muy pobladas de la periferia (San Cristóbal, Kennedy, Engativá) y son más altas en las localidades poco pobladas del centro (Santa Fe, Candelaria, Los Mártires). Esto es debido al hecho que el volumen de población varía de manera significativa de una localidad a otra, y

⁴ “Es el error asociado a la aleatoriedad del sorteo de la muestra, ya que, a partir de una muestra, en el cálculo de una estadística, se obtiene un valor de entre todos los posibles valores de la distribución del muestreo de esta estadística” (Vaillant, 2005 : 8; traducción personal)

que el número de encuestas por localidad es más o menos equivalente (para tener la misma precisión en cada localidad)

TABLA 6: tasas de muestreo del formulario ampliado por localidad en 2005

Localidad	Número de hogares encuestados (ampliado)	Número de hogares censados (basico)	Tasa de muestreo
Usaquén	5652	132302	4,27%
Chapinero	4734	48264	9,81%
Santa Fe	7697	31567	24,38%
San Cristóbal	1987	110574	1,80%
Usme	2992	78093	3,83%
Tunjuelito	2828	52470	5,39%
Bosa	3152	132917	2,37%
Kennedy	4713	257384	1,83%
Fontibón	3964	90105	4,40%
Engativá	4469	226911	1,97%
Suba	6857	266593	2,57%
Barrios Unidos	4437	61258	7,24%
Teusaquillo	3101	49857	6,22%
Los Mártires	6276	28456	22,06%
Antonio Nariño	3921	30218	12,98%
Puente Aranda	4462	73395	6,08%
Candelaria	3338	8199	40,71%
Rafael Uribe	3865	104186	3,71%
Ciudad Bolívar	3253	144641	2,25%
TOTAL	81698	1927390	4,24%

Fuentes : procesamientos realizados a partir del censo, DANE (2005)

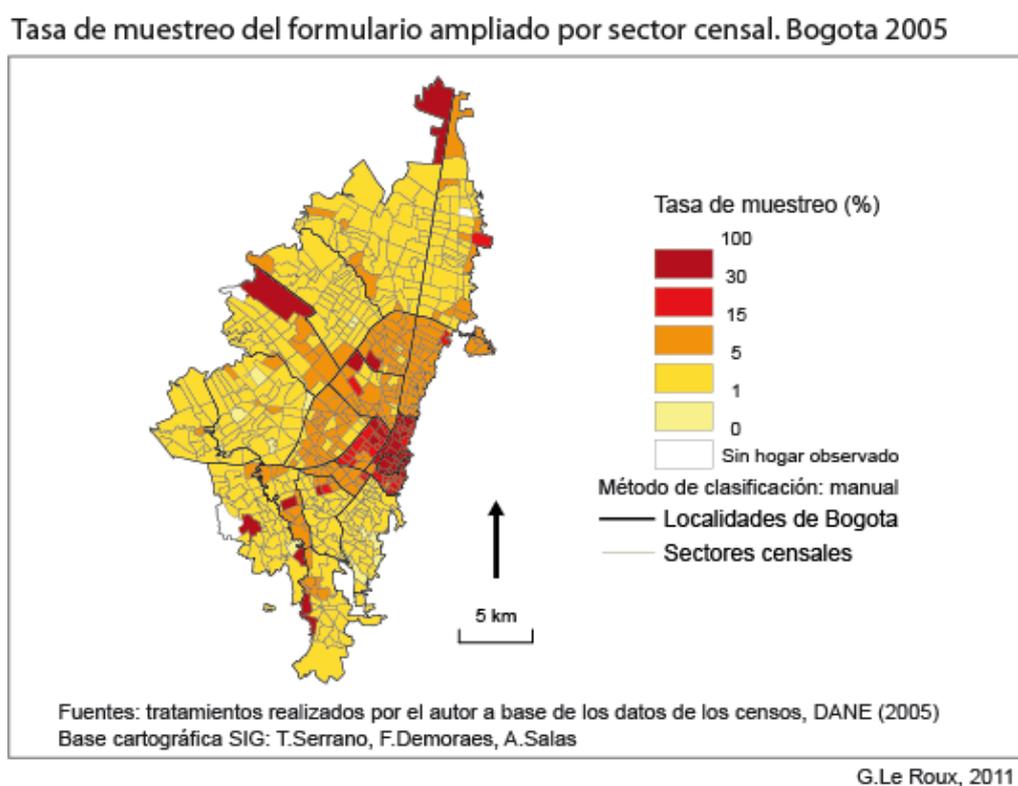
Los pesos brutos calculados son el inverso de las probabilidades de selección de los hogares en la localidad y por lo tanto mucho son más importantes en los sectores más poblados de la periferia. Estos pesos luego se ajustan (de la misma manera para todos los individuos del mismo hogar) sobre varias características de la población recogidas de manera exhaustiva (igual para las viviendas):

« Para hogar, se utilizaron como variables de calibración la edad, el sexo, la asistencia educativa, el nivel educativo máximo alcanzado y la actividad en la semana anterior al censo, de cada una de las personas que conforman el hogar; también se utilizó el total de personas del hogar. Para la temática vivienda, se utilizaron las variables: tipo de vivienda, tenencia de acueducto, de alcantarillado, de energía eléctrica y el total de hogares en la vivienda » (DANE, 2009, p.54).

Esta calibración permite obtener estimaciones bastante fiables por localidad. Sin embargo, el diseño de muestreo no garantiza una representatividad espacial para un análisis a una escala

más fina: en numerosas localidades periféricas, las tasas de muestreo son muy bajas y es posible que algunos sectores con características particulares (población o viviendas) no estén bien cubiertos por la encuesta. El cálculo de las tasas de muestreo por sectores presenta un desequilibrio entre el centro y en las periferias, pero garantiza en casi todos los sectores una tasa de muestreo de al menos 1% (FIGURA 5). Los sectores de las localidades centrales y pericentrales parecen bien cubiertos por la encuesta y se puede pensar que un análisis infra-localidad es posible sobre las variables del formulario ampliado.

FIGURA 5



3.2 Validez de un análisis infra-localidad y precisión de las estimaciones

La tasa de muestreo no afecta la precisión de las estimaciones, excepto en el caso específico de los barrios centrales, donde son especialmente altas.

A fin de aclarar la validez de un análisis infra-localidad, se evalúa la precisión de las

estimaciones que se producirían al nivel de los sectores censales a partir del tamaño de la muestra por sector. Para evaluar la precisión, evaluamos dos indicadores:

(1) el error de muestreo que surge de una encuesta en cada sector sobre una característica ficticia de los hogares; decidimos hacer este cálculo para una característica ficticia presente en el 10% y una característica ficticia presente en el 20% de los hogares

(2) la diferencia entre el tamaño de la muestra por sector y el número teórico necesario para obtener una estimación de una característica de la población con un error de muestreo predefinido; se elige como característica los individuos menores de 15 años (cuyos datos son recogidos en el cuestionario básico) y se calculan los tamaños de muestra necesarios para un error de muestreo de 10% y para un error de muestreo de 20%.

Estos dos indicadores se rigen por las siguientes ecuaciones:

$$e = t \sqrt{\frac{p(1-p)}{n}}$$

$$n = \frac{t^2 p(1-p)}{e^2}$$

donde e es el error de muestreo, t el coeficiente de margen deducido de la tasa de confianza deseado (t es 1,96 con una tasa de confianza de 95%), p la proporción de la población que tiene la característica estimada y n el tamaño de la muestra

Si la tasa de muestreo es mayor que 1/7 (como es el caso en el centro), se aplican las siguientes ecuaciones:

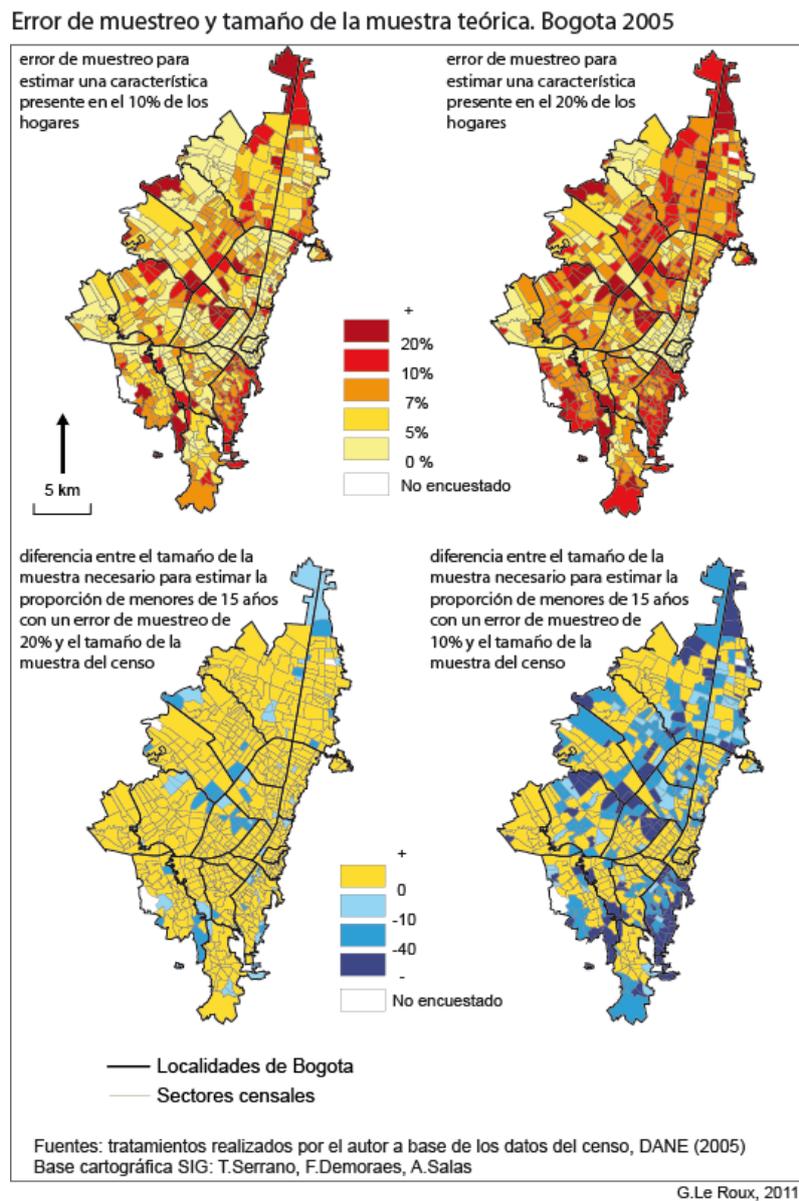
$$e = t \sqrt{\frac{p(1-p)}{n}} * \sqrt{\frac{N-n}{N-1}}$$

$$n_2 = \frac{t^2 p(1-p)}{e^2} * \frac{N-n}{N-1}$$

donde n_2 es el tamaño de la muestra teórica corregida

Los resultados muestran que los tamaños de las muestras de los sectores de las localidades centrales y de algunas localidades pericentrales permiten obtener estimaciones precisas, independientemente de la presencia de la característica estudiada en la población (FIGURA 6). En general, los errores de muestreo son raramente por encima de 10% para las características ficticias como para la característica real probada (la proporción de menores de 15 años). Los pocos casos problemáticos se encuentran dispersos en las cercanías de las periferias oeste y norte, en las periferias extremas del norte y del sur, y en los sectores de la localidad de San Cristóbal, que parecen tener muestras demasiado pequeñas para que las estimaciones sean precisas. El mapa que presenta las estimaciones de las proporciones de menores de 15 años con un error de muestreo de 20% indica que los tamaños de la muestra parecen ser adecuados en casi todos los sectores para una precisión baja.

FIGURA 6

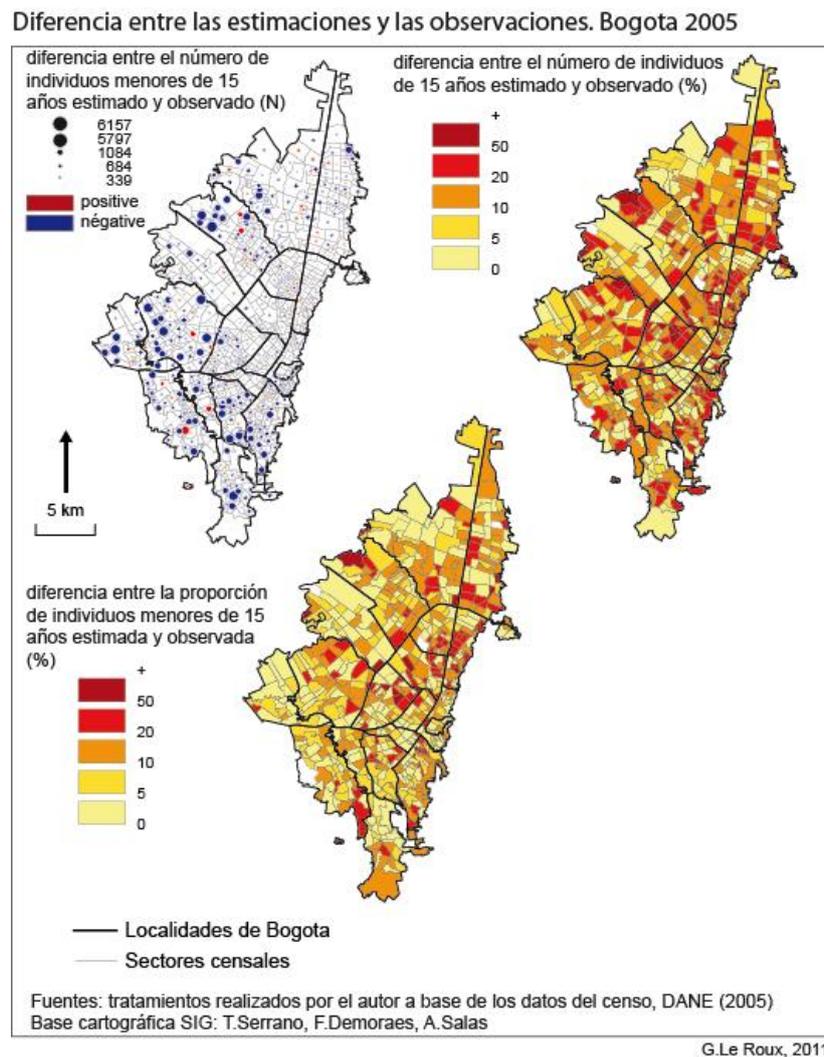


Para evaluar la calidad de las estimaciones producidas por la encuesta, se las compara con las características observadas recogidas de manera exhaustiva. Hemos escogido comparar totales (número de individuos menores de 15 años por sector) y proporciones (proporción de individuos menores de 15 años). Los pesos puestos a nuestra disposición por el DANE son los pesos calibrados sobre las características de la población por localidad. Estos pesos son problemáticos para producir estimaciones por sector, ya que no dependen de la población del sector sino de la población de la localidad. Por lo tanto, los desequilibrios en los tamaños de

población por sectores y de los tamaños de la muestra por sector en la misma localidad tendrán repercusiones en la calidad de las estimaciones. Nuevos pesos han sido producidos, iguales a la inversa de la probabilidad de selección del hogar en el sector. Estos pesos no fueron calibrados sobre las características de la población, sin embargo, proporcionan mejores estimaciones por sector que si utilizamos los pesos calibrados.

Los resultados muestran que, aunque los totales son pobremente estimados, las diferencias entre las proporciones registradas y estimadas son muy similares (FIGURA 7). Los sectores donde las diferencias son mayores se superponen a veces con los sectores ya identificados por sus pequeñas muestras, mientras que otros aparecen en el pericentro y la periferia norte. En general, cuando se considera todo el Distrito, las estimaciones de las proporciones son satisfactorias.

FIGURA 7



Estos resultados sugieren que con pesos adaptados a los sectores, utilizando el mismo procedimiento de calibración que el utilizado por el DANE, sería posible producir estimaciones confiables y razonablemente precisas por sector. El análisis anterior permite subrayar los sectores que son problemáticos por la fragilidad de sus muestras. Sin embargo, si produce resultados específicos a ciertas áreas, tales como barrios en el centro, o de una manera más amplia para mostrar las tendencias en el conjunto de Bogotá, la producción de las estimaciones a nivel de los sectores parece posible. El diseño de muestreo elegido por la encuesta del DANE no simplifica la tarea y sería interesante que en el futuro el diseño de muestreo esté estratificado a diferentes niveles geográficos inferiores a la localidad (sectores,

manzanas), para garantizar la representatividad de todo el Distrito de Bogotá, pero también para permitir análisis intraurbanos más finos.

CONCLUSIONES Y RECOMENDACIONES

En esta ponencia se pretende mostrar algunas de las dificultades que enfrentan los investigadores urbanos en la manipulación de los datos censales. El análisis de las evoluciones, y más aún el análisis a unas escalas intraurbanas, están sujetos a un problema de coherencia en las definiciones y en la manera de recoger la información. Estos problemas afectan a todos los estudios evolutivos o que utilicen diferentes fuentes de datos. Este trabajo muestra la complejidad de estudiar los cambios en general, y la dificultad de diferenciar la recogida de evoluciones reales, de la evolución en la manera de recolectar la información. Así, este estudio pone de manifiesto las diferencias en la manera de contabilizar las unidades de vivienda entre 1993 y 2005, aunque las definiciones dadas a los encuestadores sean las mismas. Un análisis de la evolución a partir del estudio de estos censos tiende a exagerar, por ejemplo, la reducción del déficit habitacional. Por otro lado, es necesario cambiar los modos de observación, las categorías de características y las definiciones para mejorar la recolección de información, y también hay que adaptarse a la realidad y a las evoluciones reales. El debate ha estado presente en Francia, por ejemplo sobre la evolución de las categorías socio-ocupacionales, que dificultan las comparaciones con las categorías anteriores, y que han sido modificadas para ser compatibles con ciertos cambios en el mercado laboral.

Por otra parte, la recogida de ciertas características de la población a través de un estudio genera el problema de no poder trabajar sobre ellas a un nivel desagregado, especialmente cuando el diseño de muestreo no es favorable como es el caso con el formulario ampliado de

2005. Así, la política de producción de información a destinación de las instituciones políticas a veces no parece tener en cuenta las dificultades introducidas para la investigación urbana.

BIBLIOGRAFÍA

- Bussi, M. & Guermond, Y. (2001). Les nouveaux espaces des recensements de population. *Revue de géographie alpine*, 89 (1), 105-117.
- Courgeau, D. (1999). Commentaires sur l'estimation de la population légale par le recensement rénové de la population. *Cybergeo : European Journal of Geography*, debate « Recensement de la population en continu : choisir le temps et l'espace ? ».
- Cuervo Ballesteros, N. (2008). *La vivienda en Bogotá. Reformas estructurales y heterogeneidad productiva en el cambio de milenio*. (Tesis de Maestría). Universidad de los Andes, Bogotá.
- DANE (1981). *XIV censo nacional de población y III de vivienda, Octubre de 1973*. Bogotá: DANE.
- DANE (1993). Instructivos para diligenciar los formularios. *Manuel del empadronador y del supervisor*. Bogotá: DANE.
- DANE (1996). *Censo 1993. Resumen nacional*. Bogotá: DANE.
- DANE (2005). *Manuel del encuestador. Censo General 2005*. Bogotá: DANE.
- DANE (2007-a). *Censo general 2005. Manuel técnico*. Bogotá: DANE.
- DANE (2007-b). *Conciliación censal 1985-2005. Colombia. Estimación de la migración 1973-2005*. Bogotá: DANE.
- DANE (2007-c). *Conciliación censal 1985-2005. Colombia. Proceso de conciliación censal 1985-2005*. Bogotá: DANE.
- DANE (2007-d). *Encuesta Calidad de Vida 2007 para Bogotá ECVB-2007*. Bogotá: DANE.
- DANE (2008). *Censo general 2005. Nivel Nacional*. Bogotá: DANE.
- DANE (2009). *Metología, Censo general 2005*. Bogotá: DANE.
- Dureau, F. & Florez, C.E. (2008). *Ficha de descripción de las fuentes de información: BOGOTA, CENSOS (1973, 1985, 1993, 2005)*. Santiago: METAL, Taller METAL, septiembre 2008.
- Florez, C.E. (2000). *Las transformaciones sociodemográficas en Colombia durante el siglo XX*. Colombia : Banco de la Republica – Tercer Mundo Editores.
- Mallarino Consuelo, U. (2007). Borrón y cuenta nueva: las estadísticas en Colombia se reinventan a sí mismas. *Universitas humanística*, 63, 91-108.
- Manrique De Llinas, H. (1990). *La población de Colombia en 1985: Estudios de evaluación de la calidad y cobertura del XV Censo Nacional de Población y IV de Vivienda*. Bogotá: DANE.
- Tabutin, D. (2006). Les systèmes de collecte des données en démographie. En Caselli, G. *et al*, *Démographie : analyse et synthèse. VIII Observation méthodes auxiliaires enseignement et recherche*. Paris : INED, 13-64.
- Toulemon, L. (2011). Individus, familles, ménages, logements: les compter, les décrire. *Travail, genre et sociétés*, 2, 47–66.
- Vaillant, J. (2005). *Initiation à la théorie de l'échantillonnage*. Recuperado de <http://monnano.weebly.com/uploads/1/6/6/3/1663287/theoechtage.pdf>